DOCUMENT RESUME

ED 038 739 **EA 002 786**

AUTHOR	Garms, Walter I.; Smith, Mark C.
TITLE	Development of a Measure of Educational Need and Its Use in a State School Support Formula. Report on Study of the New York State School Support Formula.
INSTITUTION	New York State Educational Conference Board, Albany.
REPORT NO	Staff Study-4
PUB DATE	Jun 69
NOTE	92p.
EDRS PRICE	EDRS Price MF-$0.50 HC-$4.70
DESCRIPTORS	Academic Achievement, Arithmetic, Data Analysis, Data Collection, *Educational Needs, Measurement Instruments, Predictive Measurement, Reading Ability, Research Methodology, *School Districts, *School Support, *State Aid, Statistical Analysis
IDENTIFIERS	New York

ABSTRACT

This study concerns the development of a measure of educational need and its use in a State school support formula in New York. The study assumes that public schools should operate positively to further equality of opportunity and that schools are able to affect achievement levels and rates of learning. The present State-aid formula is not sufficient because it recognizes differences in fiscal need but not in educational need. Noting that educational costs for children with special needs are relatively high compared with costs for normal children, the present formula discriminates against districts having a high proportion of special problem students. The research in this study, based on a sample of 45 school districts, suggests using the following variables: (1) The ratio of Negro and Puerto Rican students in the school, (2) the percentage of children from broken homes, (3) the average number of different schools the students have attended in the past 3 years, and (4) the average number of years of schooling the parents have completed. These measures predict 71% of the variation in educational achievement in both reading and arithmetic and, therefore, can establish educational need consistent with the stated assumptions. [Table XI on page 66 may reproduce poorly due to small print.] (LN)

EA 002 786

ED038739

REPORT ON STUDY OF THE NEW YORK STATE SCHOOL SUPPORT FORMULA

Staff Study No. 4

DEVELOPMENT OF A MEASURE OF EDUCATIONAL NEED AND ITS USE IN A STATE SCHOOL SUPPORT FORMULA

By

Walter I. Garms

and

Mark C. Smith

Teachers College, Columbia University

June, 1969

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

NEW YORK STATE EDUCATIONAL CONFERENCE BOARD • 152 WASHINGTON AVENUE • ALBANY, NEW YORK 12210

ED038739

TABLE OF CONTENTS

Preface 1

Chapter I: Introduction 2

Chapter II: Variables, Sampling, and Data Collection 19

Chapter III: Data Analysis 33

Chapter IV: Developing a State Aid Formula That Incorporates a Measure of Educational Need 53

Chapter V: Summary and Recommendations 71

Appendix A: Letter to Superintendents 81

Appendix B: Instructions for Collection of Data 83

Appendix C: Data Sheet 86

Appendix D: Reimbursement Form 87

Appendix E: Summary of Raw Data 88

LIST OF TABLES

Table I: Schools in Stratified Sample 29

Table II: Sample Returns 31

Table III: Names and Definitions of Variables Used in Analysis of 45-School Sample 34

Table IV: Summary of Data Means for Sample of 45 Schools 35

Table V: Summary of Data Ranges for Sample of 45 Schools 36

Table VI: Correlation Coefficients for 45-School Sample 39

Table VII: Proportion of Variance in R + A Explained by Different Combinations of Independent Variables 40

Table VIII: Prediction Equations for Four Sets of Variables 48

Table IX: Comparison of Average Residuals for Different Groups of Schools as Given by Various Equations 51

Table X: Comparison of Positive and Negative Residuals for Different Groups of Schools as Given by Various Equations 51

Table XI: Estimates of State Aid Distributed by Various Formulas to a Sample of Districts 66

PREFACE

This report on the development of a measure of educational need and its use in a state school support formula was sponsored and financed by the New York State Educational Conference Board. The report is one part of a major effort on the part of the Conference Board to find ways to improve the financing of public school education in New York State. Other reports have been made as part of this program by Arvid J. Burke, John W. Polley, and Robert L. Lorette.

Mark C. Smith, a doctoral student at Teachers College, has been my research assistant on this project, and has thoughtfully and capably collaborated in conceptualization, data collection, and writing. James A. Kelly, of Teachers College and The Urban Coalition, has served as a consultant to the project. The help of both of these men has been invaluable, and I gratefully acknowledge it.

This acknowledgment does not, however, relieve me of ultimate responsibility for this report, including its errors and omissions.

Walter I. Garms
Principal Investigator

CHAPTER I

INTRODUCTION

The purpose of this study is to develop a rational and practical measure of educational need and to suggest ways in which such a measure might be incorporated into the method of distributing state financial aid to school districts in New York. The study, sponsored and financed by the New York State Educational Conference Board, grew out of dissatisfaction with the present New York school support formula.

The current formula is based on local taxpaying ability and a weighted measure of daily attendance. The assumption is made that a pupil represents the same educational need in every school district regardless of his physical condition and/or social and economic environment. The inadequacy of this approach has been recognized by the state in the form of stopgap attempts to supplement the basic formula with such things as "size corrections" and "urban aid." These efforts, however, are based on questionable theoretical grounds and are inadequate to meet fully the needs generated by the concentration of large numbers of environmentally deprived children. A new approach must be taken which recognizes that some children require educational services that are more costly than those required for other children, and that the distribution of such "high cost" pupils is geographically uneven.

This Introduction focuses on four topics critical to an understanding of the thinking and procedures of this study. These topics are: the authors' concept of the role of the public school in American society; the study's working definition of educational need and how it can be measured; recent research relating socioeconomic factors with student achievement; and finally, the manner in which New York's current methods of distributing resources reflect educational need. In conclusion, the procedures followed in the study are briefly outlined.

The Role of the Public School in American Society

This study is based on two assumptions concerning the role of the school in American society. First, the public school *should* operate positively to further equality of opportunity rather than passively to perpetuate societal differences. Second, the school is *able* to affect achievement levels and rates of learning.

The idea that the free public schools should operate as a positive force for equality is central to the traditional philosophy of American education. Minority groups in America have long viewed the schools as an avenue to success and acceptance in society. Indeed, this view of the schools has been one of the prime differences between the American system of public education and the prevalent approach in most European countries.

The second assumption - that public schools can affect levels of student achievement - has come under significant questioning in recent years. Research, some of which will be cited below, has established a consistent correlation between socio-economic factors and student achievement. Some of this research has attempted to examine the relative influence of environmental factors and school factors on student achievement and has left with many readers the implication that schools make little or no difference. If this were true, a strong argument could be made for reallocating money from the schools to programs designed to improve social and economic conditions.

The best-known piece of research in this field is James Coleman's study for the Office of Education, Equality of Educational Opportunity.[1] After comparing community socio-economic factors and selected school factors with student achievement, Coleman concluded that variables measuring school effects account for little of the variance in student achievement. There are, however, theoretical and procedural weaknesses in the Coleman study which cast some doubts on this finding and the conclusions that might arise from it.

The primary difficulty stems from the fact that under current conditions in the United States, the public schools are very similar to the communities which they serve. Community socio-economic factors, the schools, and student achievement are all highly correlated with each other and it is difficult to isolate the contributions of either school factors or community factors. In his statistical treatment of the data, Coleman entered the community factors first. Once he had accounted for socio-economic differences he had accounted for most of the variation in school effects. Reanalysis of Coleman's data has shown that if the researcher takes school factors into consideration first and community factors second, the apparent effect of the school is significantly greater.[2] Samuel Bowles, who is using the Coleman data to reanalyze the relationship between school inputs and achievement has noted:

> Preliminary analysis of the computer runs which form the basis of the section of the [Coleman] Report on the effects of school resources indicate that the achievement levels of Negro students are particularly sensitive to the quality of the teaching staffs assigned to them . . . . While these results must be subjected to further scrutiny, the implication is that contrary to Coleman's conclusion, significant gains in Negro students' achievement levels can be made by directing additional resources to their education.[3]

The factor which Bowles found to be most closely associated with student verbal achievement was the teacher's score on a verbal facility test.[4]

Coleman's study is open to further question regarding his measures of school factors. The study used per-pupil expenditures by district as one measure of school input. This approach ignored important differences among schools, especially in the large cities. His use of volumes per student in the school library and the presence of science laboratories as the principal measures of school facilities is questionable. Finally, Coleman's use of total students per teacher for an entire school ignored significant variations in class size within schools.

There is a body of research which indicates that certain school characteristics do have an effect on the achievement of students. Mollenkopf and Melville found that cost of instructional support per pupil and the number of specialists on the school staff showed relatively high relationships with test scores after parental and community characteristics were controlled for.[5] In a longitudinal study with PROJECT TALENT data, Shaycroft also found that differences in schools account for significant variations in academic and vocational training.[6] It might be noted that although the conclusions differ from Coleman's, these studies necessarily faced similar difficulties in disentangling community variables from those attributed to the school.

The available evidence as to the efficacy of the schools in affecting achievement is inconsistent and beset by procedural difficulties. However, we assert our belief that schools should work to equalize opportunities and that properly financed schools can positively affect student achievement.

Educational Need

American public schools should provide all youth, regardless of socioeconomic background, with equal educational opportunities. In the past, this concept of equality of educational opportunity meant the provision of equal school facilities. The job of the school was to provide a place for learning, but the school was not held responsible for the results of the process. This approach has failed. Large segments of the population - primarily poor people and members of minority groups - are not learning the basic skills of reading, writing, and arithmetic essential for further education and for employment.

In American society today the most meaningful concept of equal educational opportunity is the opportunity for equality of educational outcome or achievement among all segments of society. This does not mean that each individual must achieve equally. Individual differences in ability, industry, and rate of learning clearly make that impossible. It does mean that consistent and significant differences in average levels of achievement between socio-economic groups must be decreased. It means that children who live in inner cities and rural districts should have the opportunity to achieve on equal levels with the more advantaged children of wealthier suburbs. The fact that this is not currently the case is well documented. A recent study in Pennsylvania found that urban districts with 25% of the average daily membership (ADM) of the state accounted for 66% of the underachievers (those with average test scores of one half grade or more below the norm). Rural districts with 26% of the ADM accounted for 26% of the low achievers. The suburban districts with 49% of the state's ADM had only 8% of the state's low achievers.[7] For the school year 1967-68, 50.8% of the third graders in New York City public schools scored below the fourth stanine (the state's definition of underachiever) on the New York state achievement test in arithmetic, whereas only 12.5% scored below the fourth stanine in the public schools in the rest of the state. Research (see below) substantiates the fact that lower socioeconomic students who tend to be concentrated in urban areas consistently and significantly achieve below students of higher socioeconomic background. These students are severely handicapped in the competition for positions in institutions of higher education and for employment opportunities.

This study defines educational need in terms of student achievement. Educational need exists wherever average achievement levels are consistently and significantly below the norm. Arguments about why the achievement in a given place is low do not deny that the need for more or better educational services exists. The schools must have the resources and initiative to make achievement approach normal levels. If educational need is defined in terms of educational achievement and if resources are applied according to some measure of this need, differences in average achievement levels among different social, economic, and racial groups will hopefully decrease and the public schools will operate in a positive manner to stimulate equal opportunity rather than perpetuating societal inequities.

The most direct measure of educational need as we have defined it would clearly be pupil achievement as indicated on test scores. Since the purpose of this study is to develop a way of distributing state aid which more accurately reflects educational need, the most obvious method would be to allocate funds in accordance with test results. There are, however, several factors which make the use of achievement scores unsuitable as a criterion for the distribution of state aid, and we have rejected this approach. Three of these factors are:

1. Low achievement may indicate an inefficient educational program yielding low return per dollar. Extra aid in this case would be rewarding inefficiency. A corollary of this problem is that state aid inversely related to achievement results could be interpreted as incentive for teachers to teach poorly, or as extra pay for a job poorly done.

2. If funds were allocated for low achievement, aid would presumably have to decrease as achievement went up, thus denying funds to effective programs.

3. The use of a standard test for distribution of state aid would raise questions regarding the validity, reliability, and cultural bias of the tests employed.

A second alternative, and the one chosen for this study is to find some measure or measures which correlate highly with student achievement. As noted above, a number of major research studies made in recent years have established a remarkably close relationship between socioeconomic factors and pupil achievement. The following section summarizes some of the more significant studies in this line of research.

Review of Major Studies of Socioeconomic Factors and Student Achievement

The types of socioeconomic factors used to examine the relationship between socioeconomic status and school achievement vary considerably, but the consistently significant correlations achieved are remarkable. Husen comments on this relationship in the summary of the International Study of Achievement in Mathematics as follows:

> The general consistency of the positive relationship between student's mathematics achievement and parental characteristics is striking. When this finding is seen in the light of the research literature, it appears that parents with higher socioeconomic characteristics do a better job of preparing their children for school (no matter what the educational system) than do parents with lower socioeconomic characteristics.[9]

Wolf and Dave's work at University of Chicago has resulted in some of the most impressive correlations between home environment and both achievement and intelligence. Using a list of 13 variables to measure individual home environments, Wolf got a correlation coefficient of $r = .76$ for student I.Q.. Using the same measure of environment, Dave found a correlation of .80 with achievement.[10]

In a study already referred to, Coleman used a list of eight variables to measure socioeconomic status of students. This list included urbanism of background, parent's education, structural integrity of the home, smallness of the family, items in the home, reading material in the home, parental interest, and parent's educational desires. Coleman's finding that these variables correlated more highly with achievement as measured by verbal ability than did school variables has already been noted.[11]

In a series of studies at the Institute of Developmental Studies, Martin Deutsch and Bert Brown divided 543 urban school children into socioeconomic strata based on "prestige ratings of occupation," "education of the main breadwinner," and "housing conditions." They found significant differences in achievement between SES levels. They also noted that Negro children at each of the three SES levels scored lower than white children and the difference increased between grades one and five.[12]

Using a scale similar to that of Deutsch and Brown, Vera John also found consistent differences in intellectual levels among students of different socioeconomic levels. Her scale was based on a combination of status of occupation, educational level of the family head and person to room ratio of the family.[13]

The International Study of Achievement in Mathematics used occupational level and level of educational attainment as two separate measures of socioeconomic status. The study concluded in part that these parental variables are significantly related to mathematics achievement in all countries studied. The tables indicated moreover that parental variables are more important in America than in most other countries.[14]

A number of studies have indicated that academic achievement and aspiration of the individual is related to the socioeconomic make-up of the student body as a whole. The classic study of the relationship of the school's social climate with achievement is that of Alan Wilson. Wilson grouped eight high schools into three socioeconomic levels on the basis of the occupational and educational background of the student body. He then correlated academic achievement and college aspiration with parental occupation, education and with the socioeconomic level of the school. The study indicated not only a high correlation between achievement and individual SES, but also that the SES of the school modified all correlations.[15]

Median family income was found to be the most significant socioeconomic variable in Burkhead's study of school achievement in Chicago and Atlanta. Burkhead initially tested five socioeconomic factors including median family income, education of parents, percentage of nonwhite population, percentage of white collar workers and unsound housing. He found that median family income accounted for a greater amount of variation in achievement than any other single school or community variable tested, although housing conditions had a high correlation in Atlanta.[16]

One of the most impressive studies of the relationship between income and success in school is Patricia Cayo Sexton's study of elementary schools in a midwestern city. Miss Sexton used average family income as an index of social class for areas served by the city's elementary schools. She compared the income level of the school to scores on the Iowa Achievement test, I.Q., and failures for grades four, six, and eight. Sexton found:

1. All schools above $7,000 income were achieving above grade level (with one exception in the eighth grade). All schools below $7,000 income were achieving below grade level.
2. Achievement test scores tended to go up as income levels go up.
3. In the fourth grade, the highest income level group was achieving two full years above the lowest income group.

She found the same relationship with I.Q. scores and with school failures. The percentage of non-promotion for the $3,000 to $5,000 level for example was 7.4%. The percentage for the $9,000 and up group was 1.2%.[17]

Francis Cornell examined the relationship of certain socioeconomic factors with achievement in his 1966 study of school finance in New York State. Cornell found high correlations between underachievement and "percentage of housing units not owner-occupied," percent of housing units not in one-unit structures, median family income of the district and percent of families with incomes under $3,000. The variable which Cornell found to be most highly correlated with underachievement was a measure of the "percentage of economically deprived children in a district." Economically deprived children were defined as children from families whose income is less than $2,000 and which are receiving Aid to Dependent Children.[18]

A number of recent research studies have attempted to go beyond the correlation of socioeconomic factors and student achievement and to examine possible causes for this relationship. Hess and Shipman have commented on the direction of this research as follows:

> The thrust of research and theory is toward conceptualizing social class as a discrete array of experiences and patterns of experience that can be examined in relation to the effects they have upon the emerging cognitive equipment of the young.[19]

Perhaps the most notable of this research centers around Bernstein's theory that language structures and conditions what the child learns and how he learns, by setting limits within which future learning takes place. Bernstein identifies two forms of communication codes or styles of verbal behavior - restricted and elaborated. By conceptualizing language as a form of social behavior his theory attempts to explain how cognitive development is affected by the verbal behavior of the home.[20] Bernstein's work has received some support in studies conducted by Hess and Shipman.[21] Others who have explored how social and economic factors effect learning include Deutsch,[22] Ausubel,[23] Strodtbeck,[24] and Bloom.[25]

The thrust of the research cited above provides support for the belief that socioeconomic factors can be found which correlate highly with achievement and which can be used as an alternative measure of educational need.

Current State Aid Programs and Educational Need

A final question essential to the background of this study is consideration of the extent to which the current method of distributing aid in New York reflects "educational need." The basic formula for distribution of state school aid in New York measures educational need in the local school district by weighted average daily attendance or, essentially, by counting pupils. Educational need as defined in this study is not taken into consideration in the basic formula. To supplement this formula, however, New York has added three different size corrections.

The results and operation of these corrections were examined thoroughly by Cornell who found that size was not a valid measure of differences in cost between districts and that the size corrections do not reflect need except to the extent that the density correction provides more aid to urban areas. Concerning the sparsity correction, Cornell concluded,

> . . . the initial form of size correction was that which allowed for presumed small class sizes and hence, higher staffing ratios in small schools. . .As it now stands, the size correction bears no relationship to cost differentials arising because of size of school or sparsity of pupil population. Moreover, such corrections for very small districts have encouraged the continuation of inefficient school district organization.[26]

After examining the two density corrections, Cornell again stated that, "Size alone is an invalid criterion or measure for determination of unusual and varying cost conditions of school districts." Furthermore, he continued, "prevalent theoretical formulations concerning cost in relation to size of school district are not supported in data from New

York school districts." Cornell argued that social and economic characteristics of the district provided sounder criteria for cost differences.[27]

In 1968-69 the state made an effort to provide some aid based on educational need as defined in this study. The attempt, called "urban aid," is in the form of providing additional money to be distributed to school districts based on need as it is measured by a combination of ADC (Aid to Dependent Children) pupils and scores on achievement tests. This program comes closer to a true measure of educational need than any of the previously cited approaches. There are a number of dangers, however, to using ADC as a measure of need. The most obvious stems from the recent Congressional decision to place a ceiling on ADC expenditures based on the 1963 level. The lists of ADC recipients are susceptible to social policy decisions at a number of levels from the United States Congress on down to the local welfare director and thus are not entirely accurate measures of need in a district. The difficulties of using test scores have already been noted.

A final aspect of the current approach to distribution of state funds in New York is the absence of any recognition of differences in need among districts for special education programs for the physically and mentally handicapped. This approach appears to be based on the assumption that such handicaps are evenly distributed throughout the state's school districts. Such an assumption is open to question. The effect of this approach may be to further handicap districts containing large numbers of poor and minority groups.

Procedures of the Study

The study of the measurement of educational need is based on five positions developed above: the schools must be a positive force in creating true equality of opportunity; the schools can affect achievement levels; the most meaningful concept of educational need is based on educational output or achievement; research indicates that socioeconomic factors can be used to predict achievement; and the current method of distributing aid in New York inadequately reflects educational need. With this background, the study has attempted to develop a useable measure of educational need for which data can easily be collected and which can be incorporated into the formula for distribution of state aid in such a way as to distribute resources more equitably in terms of educational need. The procedures of the study were as follows:

1. Identification of socioeconomic variables which are pertinent, current, easily collected, and appear to be correlated with pupil achievement.

2. Selection of a sample of schools in New York State and a subsample of students in each school.

3. Collection of socioeconomic data and achievement scores for the students in the subsamples.

4. Statistical analysis of the relation between the socioeconomic variables and achievement scores for the students and for the schools in the sample.

5. Selection of the most significant and useable variables.

6. Development of recommendations for the incorporation of the variables selected into a formula for the distribution of financial resources.

7. A test of the possible effects of these recommendations on some selected school districts.

8. The development of final recommendations.

The ensuing chapters of this report will consider the variables selected as a measure of educational need, the sample, the collection of data, data analysis, a consideration of the use of measures of educational need in the state aid formula, and recommendations for the use of our findings in the distribution of state aid.

Footnotes to Chapter I

1. James Coleman, et al., Equality of Educational Opportunity (Washington: U.S. Office of Education, 1966).

2. Marshall S. Smith, "Equality of Educational Opportunity: Comments on Bowles and Levin," The Journal of Human Resources III: 3 (Summer, 1968), 384-89.

3. Samuel S. Bowles, "Towards Equality?" Harvard Educational Review 38 (Winter, 1968), 93-94.

4. Ibid., 94.

5. William Mollenkopf and David Melville, A Study of Secondary School Characteristics as Related to Test Scores (Princeton: Educational Testing Service, 1956), as quoted in Henry S. Dyer, "School Factors and Equal Educational Opportunity," Harvard Educational Review 38 (Winter, 1968), 38-56.

6. Marion F. Shaycroft, The High School Years: Growth and Cognitive Skills (Pittsburgh: American Institutes for Research, 1967), as quoted in Dyer, op. cit.

7. Fels Institute of Local and State Government, Special Education and Fiscal Requirements of Urban School Districts in Pennsylvania (Philadelphia: Fels Institute, 1964), 9.

8. Computed from data supplied by the New York State Department of Education, Division of Testing.

9. Torsten Husén , International Study of Achievement in Mathematics (New York, 1967), 254.

10. Ravindrakumai Dave, "The Identification and Measurement of Environmental Process Variables that are Related to Educational Achievement," unpublished Ph.D. dissertation, University of Chicago, 1963. Richard Wolf, "The Identification and Measurement of Environmental Process Variables Related to Intelligence," unpublished Ph.D. dissertation, University of Chicago, 1964. See also, Robin H. Farquhar, "Home Influences on Achievement and Intelligence: An Essay Review, " Administrator's Notebook XIII (Jan., 1965).

11. Coleman, op. cit., chapter 3.

12. Martin Deutsch and Bert Brown, "Social Influences in Negro-White Intelligence Differences," The Journal of Social Issues 20 (April, 1964), 24-35. See also Gerald Lesser, Gordon Fifer, and Donald Clark, Mental Abilities of Children From Different Social

Class and Cultural Groups, monograph for Society of Research in *Child Development*, *University* of Chicago, 1965.

13. Vera John, "The Intellectual Development of Slum Children," *American Journal of Orthopsychiatry* 33 (Oct., 1963), 813-22.

14. Husén, *op.cit*.

15. Alan Wilson, "Residential Segregation of Social Classes and Aspirations of High School Boys," *American Sociological Review* 24 (Dec, 1959), 836-45.

16. Jesse Burkhead, *et al*. *Input and Output in Large City High Schools* (Syracuse: Syracuse University Press, 1967).

17. Patricia Cayo Sexton, *Education and Income: Inequalities of Opportunity in Our Public Schools* (New York: Viking Press, 1961).

18. Francis G. Cornell, "An Analysis of New York State Aid Correction," The New York State Department of Education, December, 1966. Unpublished.

19. Robert Hess and Virginia Shipman, "Early Experience and the Socialization of Cognitive Modes in Children," *Child Development* 36 (1965), 870.

20. Basil Bernstein, "Language Development and Cognition," *Educational Research* III (1961), 163-76.

21. Hess and Shepman, *op.cit*.

22. Martin Deutsch, "The Role of Social Class in Learning Development and Cognition," *American Journal of Social Issues* 20 (April, 1964), 24-35.

23. David P. Ausubel, "How Reversible Are the Cognitive and Motivational Effects of Cultural Deprivation? Implications for Teaching the Culturally Deprived Child," in Passow, Goldberg, and Tannenbaum, *Education and the Disadvantaged* (New York, 1967).

24. Fred L. Strodtbeck, "The Hidden Curriculum of the Middle Class Home," in Passow, *et al*., *op.cit*.

25. Benjamin Bloom, *Stability and Change in Human Characterics* (New York: John Wiley and Sons, Inc., 1964).

26. Cornell, *op.cit*., 94.

27. *Ibid*., 47

CHAPTER II

VARIABLES, SAMPLING AND DATA COLLECTION

In Chapter I we defined educational need in terms of achievement. Educational need is present wherever achievement is consistently and significantly below normal levels. In order to develop a measure of educational need based on socioeconomic factors which would be highly predictive of achievement it was necessary to select a number of socioeconomic variables and test their predictive power on a representative sample of schools in New York State. Chapter II describes the selection of both achievement and socioeconomic variables, the compilation of a sample, and the procedures followed in the collection of necessary data.

Selection of Variables

To test properly our approach to the measurement of educational need, it was necessary to select achievement and socioeconomic variables. Both school achievement and socioeconomic status are abstract concepts and not susceptible to direct measurement. It was therefore necessary for us to choose variables that would adequately measure these abstract concepts. Although school achievement tests measure only part of what the schools are supposed to accomplish, test scores are a practical measure of pupil achievement. The only two achievement variables which can be collected on a common basis throughout New York state are the results of the state-wide achievement tests in reading and arithmetic. These tests are given each fall in the third, sixth, and ninth grades. We recognize that these achievement tests do not measure all of the kinds of things the schools attempt to teach, but they are objective measures of some very important learning areas, and the results of them would probably correlate well with measures of achievement in other areas. As our achievement variable we have used the percentage of students in the sample schools who score below the fourth stanine on the state tests. (The fourth stanine has nothing to do with grade level. A stanine is a standard way of dividing all of those who take a test into nine groups. In any test, approximately 24% of those who take the test will fall below the 4th stanine). This measure is used by the State Department of Education and they believe it to be the most accurate way of interpreting results of the test. Those who score below the 4th stanine may be considered low achievers. Our achievement variable, therefore, is a measure of low achievement. For analysis purposes, the underachievement scores for reading and arithmetic were added together making the dependent or criterion variable for the study "percent below standard in reading plus percent below standard in arithmetic."

The selection of socioeconomic variables presented more difficulties. We were not looking for the single key environmental factor which contributes most to learning and achievement, but rather for a cluster of variables which together would serve as a plausible proxy for socioeconomic status and a reasonable measure of educational need. Since most such measures are highly correlated with each other, they are to some extent interchangeable. In selecting variables to test, it was necessary to find factors which from previous research or force of logic seem to have a relation with achievement. Our choice of variables on which to gather data could not be decided outside the context of the decision on our unit of analysis. For reasons that we believe to be sound we selected the individual school as the unit of analysis (see the later section of this Chapter entitled "Sampling Procedures" for our reasons for this choice). This choice both simplified and complicated the data gathering. It complicated it because it quickly became evident that no data which has already been gathered would suffice. Aside from the fact that U.S. Census data are almost ten years old, they are not gathered in such a way that they can be made to apply to an individual school. Other published data have the same problems. Furthermore, we find that there are no good ways to define adequately an attendance area. Busing, overlapping attendance areas, open enrollment and specialized schools obviate this. Thus we are forced into collecting data on the children who attend a school. Because of the size of the job these must be data that can readily be collected by local school personnel.

But this complication also brings its rewards. We can be confident that data gathered in this way will apply completely to the individual school, and it will be current data. We believe that such data are the only logical basis for a formula if it is to use socioeconomic data at all.

Our reading of other research studies (see Chapter I) gave us a number of possible candidates as variables to measure socioeconomic status. Some of these could be rejected out of hand as being inapplicable or impossible to gather. We were left with eleven possibilities which we examined according to the following criteria:

1. Each variable should be capable of unambiguous definition.
2. Data on each variable should be capable of being gathered currently by school clerical personnel.

3. The variable should be as free as possible of political sensitivity.
4. The variable should not be subject to influence by the school.
5. The variable should be as stable as possible.

Employing these criteria, we examined the following variables for possible inclusion in the study:

Family income. Research has indicated that family income is perhaps the best single measure of socioeconomic status. It also has been found by Burkhead, Sexton, Cornell and others to have a high correlation with student achievement (see Chapter I). The difficulty with income is that it is not obtainable for individuals without invading privacy, nor can it be easily verified. It thus violates criteria 2 and 3 above.

Occupation of family head. Father's occupation has been used in a number of studies as a simple measure of socioeconomic status. Occupations, however, are difficult to classify without a trained data collector and detailed information. For this reason, occupation could not be easily collected by school personnel.

Educational Attainment of parents. Parents' education has been shown to be positively correlated with student achievement and is a useful proxy for socioeconomic status. It can be easily collected at the time the student is registered at the school and is not politically or socially sensitive. Parents' education meets all of the criteria.

Race of ethnicity. In New York this measure would be primarily Negro and Puerto Rican. We also included in this measure a record of those children, other than Puerto Ricans, who come from homes where English is not the principal language. Ethnicity has some definition problems and some political sensitivity. It has proved to be so useful for both the schools and minority groups themselves, however, that the necessary data are currently being gathered by many schools.

Broken homes. This is a measure of whether or not the child lives with both parents. This variable may be somewhat controversial in the inner city because of the welfare implications of father absence. However, most schools routinely collect this information, and it meets the test of the other criteria.

Welfare or Aid to Dependent Children. ADC information has the advantage of being collected already by the state for the distribution of urban aid. It also correlates well with achievement. The major difficulty with this variable was noted in Chapter I: its dependence on political decisions. This objection might be lessened if welfare is only one of a number of variables used together.

Overcrowded housing. A measure of overcrowded housing could be derived by asking the number of rooms in the dwelling and the number living in the dwelling. The major difficulty with this measure is the definition of what constitutes a room, although the United State Census definition could be used. Overcrowding is indicative of socioeconomic status and generally meets the criteria established.

Substandard housing. Data on substandard housing could not be gathered by local school personnel. The state of Michigan has used as a measure the percentage of housing in a school attendance area that qualifies for urban renewal. Such a measure would only apply within cities, not statewide. Like welfare data, it would be subject to outside political influence. We have already mentioned the problems of defining a school attendance area.

Student mobility. Student mobility is a measure easily attainable from student records and is noncontroversial. It meets all of the criteria.

Population density. Population density is necessarily bound up with a school's attendance area, rather than with the students who attend the school. It was necessary to reject it on that basis.

Absenteeism. Absenteeism might provide a measure of the socioeconomic level of a school. It suffers from theoretical problems in that the school should be operating to cut down on absenteeism and it clearly does not meet criterion number four that the variable must not be subject to the influence of the school.

From this list of possible variables, six were selected for use in the study. The six which seemed to best meet the criteria established and to promise the best results are ethnic background, broken homes, welfare, parental education, overcrowded housing, and student mobility. The variable labeled ethnic background is broken down into Negro, Puerto Rican, other non-English speaking, and all others. Broken homes is simply whether two, one, or no parents live at home with the child. Welfare is a direct measure of whether the child is on welfare or not. Parents' education is the number of years of schooling of the father and of the mother (when the data were analyzed, the measure used for education was schooling of the father when present, otherwise mother's schooling). The measure of overcrowded housing is the number of rooms in the dwelling divided by the number of people residing in the dwelling. An overcrowded home is defined in this study (as it is in the U.S. Census) as one with 1.01 or more people per room. The United States Census definition is taken as the guideline of what constitutes a room. The measure of student mobility is the number of schools attended in the past three years (all students in the sample are in the fourth grade). Schools were instructed not to count natural promotional changes within the system as a change in schools. (See Appendix B for more detailed explanation of the variables used in the study).

Since the data are analyzed by school rather by individual, the actual measures aggregated by school are percent Negro, percent Puerto Rican, percent other non-English speaking, percent with one or zero parents in the home, percent on welfare, average years of fathers' schooling, percent of homes with 1.01 or more people per room, and average number of schools attended. In the test of the predictive values of socioeconomic variables, these measures are the predictor variables and the percent of students below the 4th stanine in reading and arithmetic is the criterion variable.

We anticipated the most difficulty in collecting data in the inner city schools. Therefore, in order to test the ability of the schools to collect the necessary information for the variables selected, we made a dry run in a disadvantaged elementary school in New York city. We found that the data could readily be collected by the school personnel for all of the variables.

Sampling Procedures

The next step in the development of a measure of educational need was to test the socioeconomic variables selected on a representative sample of schools in New York State. Such a test would indicate the predictive value of our variables and the ease or difficulty with which the necessary data could be collected for individuals and schools. The decision was made to use the school as the unit of analysis for a number of reasons:

1. The individual child is too small a unit of analysis. As noted in Chapter I, there will always be large variations in individual achievement because of differences in annate ability, industriousness, and rate of learning.

2. The school district is too large a unit. In most districts there are schools with a concentration of disadvantaged children and other schools with a concentration of advantaged children. When using district averages many of these schools cancel one another out, leaving a composite figure that hides real educational problems.

3. School district consolidation or decentralization would be unlikely to affect a measure based on the individual school.

4. The school is the natural unit of the educational organization and is therefore the level at which we may best look for educational improvements. As discussed earlier, we decided to collect the actual data on individual students and aggregate the data by school. In this manner we would avoid difficulties created by the lack of fixed attendance zones.

The decision to collect data on individuals and aggregate the information for analysis by school created the necessity of two-staged sampling procedure. First, a representative sample of schools had to be selected and second, a sample of students had to be chosen from within each school.

The approach used to obtain a selection of schools is a stratified, random sampling procedure. A stratified procedure was followed because we felt that any consideration of New York State finance had to include consideration of New York City. New York City plays a predominant role in state finance as well as education and a purely random sample ran the risk of under- or over-representing the City. Furthermore, we wished to examine the effect of any results on the different strata within the state as well as on the state as a whole. Within the strata, a random procedure was followed to assure that the schools selected are representative of the stratum. For the purposes of the study, school districts of the state were stratified into New York City, other large cities, medium cities, suburbs, and rural. The definitions of these are as follows:

1. New York City - self explanatory.
2. Other large cities - cities of over 100,000 population based on the 1960 U.S. Census.
3. Medium cities - cities of 50,000 to 100,000 population based on the 1960 U.S. Census.
4. Suburbs - the Research Department of the State Department of Education has categorized the remaining districts in the state as "small cities" (less than 50,000), "large rural districts" and "small rural districts." The suburban stratum consists of all districts which are categorized as small cities or large rural districts and which are located in a Standard Metropolitan Statistical Area.
5. Rural - all districts designated as small cities or large rural districts which are not located in a Standard Metropolitan Statistical Area, plus all small rural districts wherever located.

It is recognized that the classification of schools into these strata is arbitrary, but there is no method yet discovered which is not arbitrary. This classification is logical and results in only minor inequities of the sort that any and all classifications encounter.

For the selection of a representative sample of schools for each stratum, the universe from which the sample was to be chosen was narrowed by our decision to use the New York State Achievement tests as the dependent variable. The tests are given only in the third, sixth, and ninth grade and had not been given in all parts of the state for the current school year at the time the data was to be collected. This meant that the sample would have to consist of schools with fourth, seventh, or tenth grades in order to contain students who had taken the test the previous year. Because the chance of continuity in one school seemed greater for third and fourth grades, we decided to focus our efforts on students in the current fourth grade who had taken the third grade version of the test in the school year 1967-68. Finally, we decided on a subsample of 20 students from each school. To ensure the presence of 20 students in the fourth grade we wanted schools that had 25 or more students enrolled in the third grade during the 1967-68 school year. The universe for our sample of schools within each stratum consisted of those public schools which have both a third and a fourth grade and which had at least 25 students enrolled in the third grade during the school year 1967-68. The data would be collected for current fourth graders, but the achievement test results would be from their third grade year.

In order to select a sample of schools in each stratum proportional to the total student population of that stratum, the total number of public school pupils in the stratum was determined and from that the percentage that the student population of the stratum bore to the overall public student population of the state. By then taking a sample number of schools from each stratum proportional to the total student population of the stratum and twenty students from each school, the result would be a number of schools and a number of students in each stratum proportional to the total number of students in the stratum.

We decided on a sample size of 75 schools with an aim of getting at least 50 useable returns. The total number of students in each stratum, the percentage of the total and the indicated number of schools to be selected were as follows:

TABLE I

Stratum	Student Population	%	Number of Schools
New York City	1,112,501	33.3	25
Other large cities	226,603	6.8	5
Medium cities	46,595	2.9	2
Suburbs	1,312,656	39.3	29
Rural	592,690	17.7	13

Because the stratum "medium cities" contained less than 3% of the student population and would have a sample of merely two schools, we decided to combine this stratum with "other large cities." This meant that there are four final strata. They are, New York City, Other Cities, Suburbs, and Rural. Still, the second stratum was so small (seven schools) that it appeared little could be said with confidence about this stratum. Accordingly, we doubled the sample size (to 15) for that stratum. This action was taken into account in the analysis of the data. The final size of the stratified sample used in the study is 82 schools.

The selection of the sample schools within each stratum was accomplished by a simple random procedure. A list of each eligible elementary school in the stratum was compiled and the schools numbered consecutively. The appropriate number of schools was selected using a table of random numbers.

A simpler procedure was followed in selecting the pupils within the sample schools since the actual selection would have to be done by employees in the schools (see "Data Collection" below). From last year's third grade enrollment an estimated number in the current fourth grade was determined. From this number the school was requested to select every fifth pupil (or fourth or second or whatever number was necessary) to provide a sample of not less than 25 nor more than 40 pupils. The first 20 of these on which full data could be obtained would be the sample from that school. Although this procedure meant a slight digression from randomness, the restraints of time, manpower, and money made it necessary.

Data Collection

The data to be collected for each of the twenty fourth-grade students in the 82 sample schools was:

1. State reading and arithmetic scores from the third grade.
2. The student's race or ethnic status.
3. How many parents live with the child.
4. Whether the student is on welfare.
5. The number of years of schooling of his parents.
6. The number of rooms in the student's dwelling.
7. The number of people who live in the student's dwelling.
8. The number of schools the student has attended over the past three years.

This data was to be collected by the schools in the sample from the student's record cards, the student himself, and the parents.

A letter was sent to the Superintendent of schools in each district containing a sample school. (See Appendix A). With the Superintendent's approval, the letter was to be passed on to the Principal of the sample school, along with detailed instructions for collection of the data, a form for recording the data, and a reimbursement form for reimbursing the individual designated by him to collect the data. (Appendices B, C, and D). For New York City a suggested form was included to be sent home to parents and returned to the school. Telephone follow up contacts were made with those principals who were slow returning data.

Sample Returns

Table II summarizes the response of sample schools by stratum.

TABLE II

Stratum	Number	Complete Returns	Percentage Return
New York City	25	10	40.0%
Other cities	15	8	53.3%
Suburbs	29	16	56.9%
Rural	13	11	84.6%
totals	82	45	54.9%

The final sample was composed of 45 schools. Because of the higher rate of return from suburban and rural schools, the composition of the final sample does not reflect the proportion of total students in the four strata.

The major difficulties in data collection stemmed from the necessity of collecting data through the mail and the lack of direct contact. Reliance on letters and telephone calls was necessary because of the scope of the sample and the limitations of time and money. Less than 8% of the sample indicated an unwillingness to collect the data because of its semi-controversial nature. There is little question that many districts failed to see any direct benefit from such a study and are severely strapped for manpower and time at the school level.

Considerable difficulty was encountered in New York City in obtaining the scores from the state achievement tests. Because the city schools use the Metropolitan Achievement Test for measuring achievement and for placement purposes, many of the schools in the sample had no record of the results of the state test. The school either did not receive the scores from the central office or had failed to record them on the individual student's permanent record card. For the schools which compiled all of the data except state test scores we were able to obtain the scores from the central office files, but it is possible that this problem contributed to the relatively higher rate of nonresponse in the city schools. Because we did not get a higher rate of return on our sample, we are not in a position to claim that it is a true random sample, and to apply the usual tests of significance to the result. But as the following chapter on the data analysis will indicate, we have reason to believe that the sample is substantially representative and that our results are valid for the purposes to which we put them. We do not believe that the problems we encountered would apply to state-mandated data collection by all schools.

CHAPTER III

DATA ANALYSIS

In this chapter we first examine the extent to which we may place faith in our sample as being representative of schools in New York State. Next we discuss the results of using data on our variables to predict school achievement. Finally, we examine the residuals of our prediction equations by stratum to see how adequately our variables predict educational achievement in New York City, other cities, suburbs, and rural areas.

Representativeness of the Sample

We shall have frequent occasion in this chapter to refer to the variables. For convenience, we have abbreviated the names of these variables, as is shown in Table III.

Appendix E contains the raw data for the 45 schools in our sample. Table IV gives a summary of the data means for the schools in the various strata and for the entire sample of 45 schools. Table V gives the ranges of the variables in these same categories. From this information some inferences may be made about the representativeness of the sample. We were at first concerned, when we did not get back all of our questionnaires, that a process of self-selection might occurring, with the schools serving a higher socioeconomic clientele returning data at a much higher rate than those serving the lower socioeconomic groups. However, Tables IV and V indicate that we have gotten a rather representative sample. The range in each of the variables is surprisingly large, and the means are, in general, the sort that one would expect.

TABLE III

NAMES AND DEFINITIONS OF VARIABLES USED IN ANALYSIS OF 45-SCHOOL SAMPLE

READ	Percent of pupils scoring below 4th stanine in state reading achievement test
ARITH	Percent of pupils scoring below 4th stanine in state arithmetic achievement test
R + A	READ + ARITH
N	Percent of pupils who are Negro
PR	Percent of pupils who are Puerto Rican
F	Percent of pupils (other than Puerto Rican) coming from homes where the main language spoken is not English.
BrHo	Percent of pupils from broken homes (where one or both parents are missing)
Wlfr	Percent of pupils whose family is receiving Aid to Dependent Children
O-C	Percent of pupils living in overcrowded housing (where there is more than one occupant per room in the dwelling)
MOB	Mobility (average number of schools attended by pupils during last three years)
Sch-F	Average years of schooling of pupils' fathers (when present in the home)
Sch-M	Average years of schooling of pupils' mothers (when present in the home)
SCH	Average years of schooling of father when present, otherwise mother

TABLE IV

SUMMARY OF DATA MEANS FOR SAMPLE OF 45 SCHOOLS

	10 Schools in New York City	8 Schools in Other Large Cities	16 Suburban Schools	11 Rural Schools	All 45 Schools
READ	30.4%	41.4%	17.3%	25.1%	25.4%
ARITH	32.9%	24.7%	3.8%	8.7%	14.2%
N	23.5%	27.2%	0.3%	3.7%	10.7%
PR	21.1%	7.4%	0.3%	2.7%	4.9%
F	0.5%	0.0%	1.3%	0.0%	0.6%
BrHo	25.9%	21.9%	7.6%	3.8%	12.1%
Wlfr	22.2%	17.3%	1.6%	2.3%	8.0%
O-C	34.6%	26.0%	12.9%	16.8%	21.2%
MOB	1.45	1.37	1.19	1.22	1.29
Sch-F	11.1	10.7	12.9	11.8	11.9
Sch-M	10.9	10.8	12.3	11.7	11.6
SCH	11.0	10.9	12.9	11.8	11.9

TABLE V

SUMMARY OF DATA RANGES FOR SAMPLE OF 45 SCHOOLS

	10 Schools in New York City	8 Schools in Other Large Cities	16 Suburban Schools	11 Rural Schools	All 45 Schools
READ	0.0 - 61.1	5.0 - 60.0	0.0 - 40.0	0.0 - 45.0	0.0 - 61.1
ARITH	0.5 - 61.1	5.0 - 35.0	0.0 - 20.0	0.0 - 29.4	0.0 - 61.1
N	0.0 - 94.7	0.0 - 85.0	0.0 - 5.3	0.0 - 35.0	0.0 - 94.7
PR	0.0 - 55.6	0.0 - 5.9	0.0 - 5.0	0.0 - 29.4	0.0 - 55.6
F	0.0 - 5.0	0.0 - 0.0	0.0 - 10.0	0.0 - 0.0	0.0 - 10.0
BrHo	10.0 - 66.7	10.0 - 50.0	0.0 - 23.5	0.0 - 15.0	0.0 - 66.7
Wlfr	0.0 - 56.2	0.0 - 40.0	0.0 - 10.0	0.0 - 10.0	0.0 - 56.2
O-C	5.0 - 70.6	10.0 - 55.0	0.0 - 45.0	5.0 - 50.0	0.0 - 70.6
MOB	1.05 - 2.16	1.20 - 1.60	1.05 - 1.50	1.00 - 1.65	1.00 - 2.16
Sch-F	7.9 - 13.6	8.6 - 13.2	9.9 - 15.6	9.9 - 14.6	7.9 - 15.6
Sch-M	7.3 - 13.7	8.6 - 12.7	10.4 - 16.2	10.0 - 13.5	7.3 - 16.2
SCH	7.9 - 13.6	8.6 - 13.2	9.9 - 15.6	9.9 - 14.6	7.9 - 15.6

Thus, while we do not claim that this is a true random sample, we have reason to believe that the results of our statistical tests are a reasonable reflection of reality. Because the sample is not a true random one we are unable to give statistical confidence limits for our results. Instead, we present the results as being worthwhile because they show that it is possible to collect socioeconomic data that are sufficiently noncontroversial to be gathered by school personnel but that are powerful enough in their predictive powers for use in a state aid formula.

The Prediction Equations

The statistical technique used in this investigation is that of multiple regression. This technique recognizes that in many things there are several influences simultaneously at work. Alternatively, while there may be only one influence at work there may be no way to measure that influence directly. In such a case, several variables that are believed to be strongly related to the influence are measured and these variables together are used as a substitute for the unmeasureable influence. The multiple regression technique defines a prediction equation using the weighted combination of the predictor variables that will best predict a criterion variable. In the case of our study, it is our belief (see Chapter One) that academic achievement is strongly influenced by a factor called socioeconomic status (or that there is a group of influences operating that may be jointly called socioeconomic status). We wish to attempt to predict, as best we can, school achievement from some knowledge of socioeconomic status.

Our problem in this study was to find variables that could be gathered by school personnel yet would have sufficient predictive power. In looking at these variables, one should not think only in terms of how that particular variable might directly affect school achievement, as for instance in rationalizing that a child who lives in overcrowded housing finds it difficult to study. One should also look upon each of the variables as a substitute measure for socioeconomic status. Looked at in this way, then, these predictor variables are to some extent interchangeable, and our problem is to find the combination which seems best to use considering convenience, political sensitivity, stability, and predictive power. Table VI shows correlations of all of the variables, indicating that the predictor variables (with the exception of F, the percentage of non-Puerto Rican students whose principal language at home is not English) are all rather highly correlated with each other, and thus tend to measure the same thing.

Table VII shows the results of trying most of the possible combinations of one or more predictor variables in an attempt to see how much of the variation in the criterion variable R + A (percent below standard in reading plus percent below standard in arithmetic) could be predicted. An "x" or an "s" in a column indicates that that column's variable was used in the prediction equation.* The "x" indicates that the variable acts like a normal variable in the equation; the "s" that the variable acts like a suppression variable (the suppression variable is explained

---

*In order to avoid confusion, we have consistently used the word "equation" to refer to an algebraic expression which predicts school achievement using socioeconomic variables. We have used the word "formula" to refer to a state school aid formula, which may or may not contain a prediction equation as a part of it.

TABLE VI

CORRELATION COEFFICIENTS FOR 45-SCHOOL SAMPLE

	READ	ARITH	R + A	N	PR	N + PR	F	BrHo	Wlfr	O-C	MOB	Sch-F	Sch-M	SCH
READ	1.00	.72	.93	.52	.54	.65	-.24	.58	.65	.46	.43	-.61	-.58	-.60
ARITH	.72	1.00	.92	.54	.81	.79	-.18	.69	.76	.56	.50	-.62	-.60	-.64
R + A	.93	.92	1.00	.57	.73	.77	-.22	.68	.76	.54	.50	-.66	-.64	-.67
N	.52	.54	.57	1.00	.30	.89	-.13	.70	.70	.56	.35	-.54	-.52	-.52
PR	.54	.81	.73	.30	1.00	.70	-.10	.56	.62	.52	.48	-.51	-.52	-.55
N + PR	.65	.79	.77	.89	.70	1.00	-.15	.79	.82	.67	.49	-.65	-.64	-.65
F	-.24	-.18	-.22	-.13	-.10	-.15	1.00	-.00	-.09	-.11	-.04	.15	.13	.16
BrHo	.58	.69	.68	.70	.56	.79	-.00	1.00	.88	.52	.37	-.58	-.62	-.58
Wlfr	.65	.76	.76	.70	.62	.82	-.09	.88	1.00	.65	.45	-.63	-.62	-.65
O-C	.46	.56	.54	.56	.52	.67	-.11	.52	.65	1.00	.33	-.71	-.62	-.72
MOB	.43	.50	.50	.35	.48	.49	-.04	.37	.45	.33	1.00	-.21	-.21	-.26
Sch-F	-.61	-.62	-.66	-.54	-.51	-.65	.15	-.58	-.63	-.71	-.21	1.00	.93	.99
Sch-M	-.58	-.60	-.64	-.52	-.52	-.64	.13	-.62	-.62	-.62	-.21	.93	1.00	.92
SCH	-.60	-.64	-.67	-.52	-.55	-.65	.16	-.58	-.65	-.72	-.26	.99	.92	1.00

TABLE VII

PROPORTION OF VARIANCE IN R + A EXPLAINED BY DIFFERENT COMBINATIONS OF INDEPENDENT VARIABLES

Number of Variables in Equation	Run No.	N and PR*	N + PR	F	BrHo	Wlfr	O-C	MOB	SCH	Prop of Variance Explained
Seven	1	x		x	s	x	s	x	x	.75
	2		x	x	s	x	s	x	x	.71
Six	3	x			s	x	s	x	x	.73
	4		x		s	x	s	x	x	.70
Five	5	x			s	x	s	x		.71
	6	x			s	x	s		x	.73
	7	x			x	x		x	x	.72
	8	x				x	s	x	x	.73
	9				x	x	s	x	x	.66
Four	10	x			s	x	s			.70
	11	x			s	x		x		.70
	12	x			s	x			x	.72
	13	x			x		x	x		.69
	14	x			x		s		x	.71
	15	x			x			x	x	.71
	16	x				x	s	x		.70
	17	x				x	s		x	.73
	18	x				x		x	x	.72
	19	x					s	x	x	.71
	20				x	x	x	x		.61
	21				x	x	s		x	.63
	22				x	x		x	x	.66
	23				x		s	x	x	.64
	24					x	s	x	x	.66

TABLE VII (CONT.)

Number of Variables in Equation	Run No.	N and PR*	N + PR	F	Br-fo	Wl-fr	O-C	MOB	SCH	Prop of Variance Explained
Three	25	x			s	x				.70
	26		x		s	x				.64
	27	x			x		x			.69
	28		x		x		x			.61
	29	x			x			x		.69
	30		x		x			x		.63
	31	x			x				x	.70
	32		x		x				x	.65
	33	x				x	s			.70
	34		x			x	s			.64
	35	x				x		x		.70
	36		x			x		x		.65
	37	x				x			x	.72
	38		x			x			x	.67
	39	x					x	x		.67
	40		x				x	x		.62
	41	x					s		x	.70
	42		x				s		x	.65
	43	x						x	x	.70
	44		x					x	x	.67
	45				x	x	x			.58
	46				x	x		x		.60
	47				x	x			x	.62
	48				x		x	x		.57
	49				x		x		x	.58
	50				x			x	x	.64
	51					x	x	x		.61
	52					x	x		x	.63
	53					x		x	x	.66
	54						x	x	x	.56
Two	55	x			x					.68
	56		x		x					.61
	57	x				x				.70
	58		x			x				.64
	59	x					x			.67
	60		x				x			.60
	61	x						x		.67
	62		x					x		.61

TABLE VII (CONCL.)

Number of Variables in Equation	Run No.	N and PR*	N + PR	F	BrHo	Wlfr	O-C	MOB	SCH	Prop of Variance Explained
	63	x							x	.69
	64		x						x	.64
	65				x	x				.57
	66				x		x			.51
	67				x			x		.54
	68				x				x	.57
	69					x	x			.58
	70					x		x		.60
	71					x			x	.62
	72						x	x		.41
	73						x		x	.45
	74							x	x	.56
One	75	x								.67
	76		x							.60
	77			x						.05
	78				x					.47
	79					x				.57
	80						x			.30
	81							x		.25
	82								x	.44

---

* N and PR indicates that although these are conceptually one variable, they have been separately entered in the regression equation, and have separate regression weights. N + PR indicates that the data values for N and PR have been added together before entering them into the regression equation, so that it is in fact as well as concept a single variable, with a single regression weight.

in a later paragraph). The last column indicates the proportion of the variation in R + A that is explained by the predictor variables used in the equation. Note that in some equations N (percent Negro in the school) and PR (percent Puerto Rican in the school) are treated as separate variables, while in other equations they are added together to form a single variable. The equations are presented in the Table starting with those using the largest number of variables and concluding with those using only a single variable.

Run 1 of Table VII indicates the remarkable predictive power of the variables we have used. Using these variables, we can predict 75% of the variation in school achievement (as measured by R + A) without knowing anything about the instructional programs of the schools. It is this result, confirming as it does other studies of the close association between socioeconomic status and school achievement, that gives us confidence in our sample and in the variables we have chosen.

In Run 2 of Table VII we have used the combination variable N + PR instead of using N and PR separately. The result is a small but significant loss in predictive power. We have tried this in numerous other equations, as shown in the Table, with a consistent loss of predictive power. It appears, then, that any predictive equation that uses N and PR should use them as separate variables, each with its individual weight, rather than as a combination variable.

If we can predict 75% of the variation in school achievement with these variables, why not just use them all in a prediction equation? The answer, of course, lies in economy of effort. We need an equation that will do an excellent job of prediction, yet which will use no more variables than necessary. There are also some other criteria to be applied. The following is the complete list of criteria used in choosing the particular set of variables that we recommend:

1. The variables used should, as much as possible, meet the five criteria given in Chapter II.

2. The equation should be a powerful predictor of school achievement using a relatively small set of variables.

3. The equation should use at least three variables in order to diminish the chance that the variables used will not apply well to all districts.

4. The equation should not contain a variable that is acting as a suppressor variable (see the later discussion of this).

The five criteria for variables, given in Chapter II, were used to evaluate all of the variables that were considered in our initial planning for the project. Let us briefly review how the variables on which we actually collected data meet these criteria:

1. N (percentage of Negroes in the school) has some difficulties of definition and some political sensitivity. However, data on this variable are currently gathered by the New York schools, so we are not breaking new ground. The variable meets the other criteria adequately, and we decided to retain it for consideration in an equation.

2. PR (percentage of Puerto Rican students in the school) has the same strengths and weaknesses as a variable as has N. We also decided to retain it for consideration.

3. F (percentage of students, except Puerto Ricans, from homes where the principal language is not English) meets all of the criteria except

that of stability. Because there is such a small number of these students (less than 1% in our sample) the correlations of this variable with the other variables are extremely unstable. We are forced to reject this as a variable.

4. BrHo (the percentage of children from broken homes) meets all of the criteria, and was retained for consideration.

5. Wlfr (the percentage of children on welfare) does not meet the criterion of stability because it is subject to outside political influences. Although it meets the other criteria well, it should be rejected on stability grounds. We gathered data and tested it as a variable principally because it is already in use in the "urban aid" part of the state aid formula and we felt it was important to check its value as a predictor against the other variables we chose.

6. O-C (the percentage of pupils living in overcrowded housing) suffers from some problems in definition, and this makes it difficult for school personnel to gather accurate data. Also, in some areas there is some political sensitivity attached to this measure. We gathered data on it because preliminary investigation made it appear promising as a variable. We found that its predictive power was insufficient to offset the problems of data collection, and we rejected it in the final analysis.

7. MOB (the average number of schools attended by the pupils in the last three years) meets all of the criteria well, and was retained for consideration.

8. SCH (the average number of years of schooling of the father if present, otherwise the mother) meets all of the criteria well. It was retained for consideration.

Thus we are left with the following variables as the most likely candidates for an equation: N and PR, BrHo, MOB, and SCH. Let us examine the possible equations containing these variables, and apply criteria 2, 3, and 4, listed earlier, which applied to equations.

The best single predictor is N and PR (remember that we are thinking of N and PR as a single variable even though we may weight N and PR differently). With it we can predict 67% of the variation in school achievement (see Run No. 75). With seven variables we can predict only 8% more than that (see Run No. 1). Yet we cannot use only N and PR in our equation because of criterion 3, which suggests that one or two variables are insufficient.

The best combination of two of our selected variables is N and PR plus SCH, predicting 69% of the variation in achievement (see Run No. 63). We could predict 70% using Wlfr as one of the two variables (see Run No. 57), but we have rejected Wlfr for its instability. It is gratifying that there are other variables approximately as powerful as Wlfr that do not share its problems.

The best combination of three of our selected variables is either N and PR, MOB, and SCH (see Run No. 43), or N and PR, BrHo, and SCH (see Run No. 31). Both predict 70% of the variation in achievement. Using Wlfr as one of the variables would only increase this prediction 2% (see Run No. 37). A third possibility is N and PR, BrHo, and MOB, predicting 69% of the variation in achievement (see Run No. 29). All three of these are very close together in predictive power, and all deserve strong consideration for use as our predictive equation. All of them meet the three criteria for such an equation.

The equation which uses all four of the variables predicts 71% of the variation in achievement (see Run No. 15). It also meets all of our criteria for an equation.

Perusal of the Table indicates that there is not much to be gained by using any of the variables we had rejected on other grounds. In addition, in some of the combinations one or more of the variables acts as a suppression variable. A suppression variable is one that has a positive correlation with the criterion variable, but because of being highly correlated with one of the other predictor variables acquires in a particular equation a negative regression weight. There is nothing wrong with this; the equation is still a valid one. But use of a suppression variable in an equation would be hard for most people to understand. As a result, we have ruled out the use of any equation which contains such suppression variables. We are fortunate that those variables which we found acceptable on other bases do not tend to act as suppression variables.

We are thus left with four equations which meet all of our criteria:

1. N and PR, BrHo, MOB, and SCH, predicting 71%.
2. N and PR, MOB, and SCH, predicting 70%.
3. N and PR, BrHo, and SCH, predicting 70%.
4. N and PR, BrHo, and MOB, predicting 69%.

All of them are very close to each other in predictive power, and considering the limitations of the sample there is probably no significant difference among them. Table VIII gives the actual prediction formulas for each of these equations.

TABLE VIII

PREDICTION EQUATIONS FOR FOUR SETS OF VARIABLES

Equation 1:

$$R + A = 0.25\ N + 0.89\ PR + 0.29\ BrHo + 16.34\ MOB - 3.92\ SCH + 53.28$$

Equation 2:

$$R + A = 0.37\ N + 1.01\ PR + 15.89\ MOB - 4.20\ SCH + 59.20$$

Equation 3:

$$R + A = 0.31\ N + 1.03\ PR + 0.28\ BrHo - 3.59\ SCH + 69.22$$

Equation 4:

$$R + A = 0.36\ N + 1.10\ PR + 0.36\ BrHo + 13.01\ MOB + 7.89$$

Analysis of Residuals

It is important to know the extent to which these equations predict accurately for each of our strata as well as for the statewide sample. One way to do this is to examine the residuals for our sample. A residual is the difference between the actual value of the criterion variable for a particular school and the value that is predicted by the equation. For example, the prediction equation containing N and PR, MOB, and SCH is as follows:

$$R + A = 0.37 + 1.01\ PR + 15.89\ MOB - 4.20\ SCH + 59.20$$

Suppose a particular school had the following data:

READ = 40 (that is, 40% of the pupils are below standard in reading)

ARITH = 47 (47% of the pupils are below standard in arithmetic)

N = 39 (39% of the students are Negro)

PR = 25 (25% of the students are Puerto Rican)

MOB = 2.0 (on the average, pupils have attended two schools in the last three years)

SCH = 9.5 (the pupils' fathers, or their mothers where the fathers were not present in the home, had attended 9.5 years of school on the average).

Then the actual value of the criterion variable is

$$R + A = READ + ARITH = 40 + 47 = 87$$

The predicted value is

$$\begin{aligned} R + A &= (0.37 \times 39) + (1.01 \times 25) + (15.89 \times 2.0) - (4.20 \times 9.5) + 59.20 \\ &= 14.43 + 25.25 + 31.78 - 39.90 + 59.20 \\ &= 90.76 \end{aligned}$$

The residual is the true value minus the predicted value:

Residual = 87.00 - 90.76 = -3.76

For the entire sample of 45 schools, the statistical procedure insures that the sum of negative residuals will equal the sum of the positive residuals. However, we may properly ask how the various possible equations compare in predicting achievement for the schools in each of our strata. We have done this for each of the four equations we have found acceptable, with results shown in Tables IX and X. Table IX shows, for our sample, the average amount of the residual for each of the strata. Thus, the first equation, involving all of the variables overestimates the underachievement in the school by 2.77% in the 10 sample schools in New York City (a negative residual indicates that the equation overestimates). It underestimates the value for the 8 schools in other large cities by 7.56%. It overestimates by 3.81% in the suburban districts, and underestimates by 2.57% in the rural districts. Inspection of the Table indicates that on the basis of this criterion, the second equation, involving N and PR, MOB, and SCH, does the best job, although the results for all four are rather similar.

However, the average residual may be thrown off quite a bit by a single school when there are a relatively few schools in the sample. This is true of the sample of 8 schools from "other large cities," for there is one school in that sample in which achievement is much worse than would be predicted by our variables. The actual value of R + A for the school is 80, but our variables predict it to be only about 30. It is possible that this is a school with an unusually ineffective program,

TABLE IX

COMPARISON OF AVERAGE RESIDUALS FOR DIFFERENT GROUPS OF SCHOOLS AS GIVEN BY VARIOUS EQUATIONS

Formula Variables	10 Schools in New York City	8 Schools in Other Large Cities	16 Schools in Small Cities and Suburbs	11 Rural Schools
N and PR, BrHo, MOB, SCH	-2.77	+7.56	-3.81	+2.57
N and PR, MOB, SCH	-2.55	+8.53	-3.36	+0.99
N and PR, BrHo, SCH	-2.56	+9.09	-4.47	+2.20
N and PR, BrHo, MOB	-4.33	+10.33	-5.45	+4.35

TABLE X

COMPARISON OF NUMBER OF POSITIVE AND NEGATIVE RESIDUALS FOR DIFFERENT GROUPS OF SCHOOLS AS GIVEN BY VARIOUS EQUATIONS

	10 Schools in New York City		8 Schools in Other Large Cities		16 Schools in Small Cities and Suburbs		11 Rural Schools	
Formula Variables	+	-	+	-	+	-	+	-
N and PR, BrHo, MOB, SCH	4	6	5	3	7	9	5	6
N and PR, MOB, SCH	4	6	5	3	7	9	5	6
Nand PR, BrHo, SCH	3	7	5	3	7	9	6	6
N and PR, BrHo, MOB	4	6	5	3	3	13	7	4

but we are not here to pass judgement on the reasons for the discrepancy. Sufficient to say that this single school affects drastically the average residual for its stratum. A different way of looking at residuals, that eliminates this effect of a single large residual, is shown in Table X, where the number of positive and negative residuals is given for each stratum. The only prediction equation that does poorly is number 4, for in this one 13 out of the sixteen suburban schools have negative residuals.

We suspect that a larger sample would show no really significant differences among any of these four equations in terms of predictive power or in treatment of schools in different strata. It therefore seems reasonable to declare that any of these equations is acceptable. As will be seen in Chapter V, we are not recommending the immediate use of any of these equations in a state aid formula. We are instead recommending that a year be devoted to a required collection and processing of data for all schools in the state, and that a prediction equation be developed based on that solid data base rather than upon the data of a sample. A state aid formula incorporating this prediction equation would be put into use in the following year. In the next chapter we shall see how a state aid formula based upon a measure of educational need could be developed.

CHAPTER IV

DEVELOPING A STATE AID FORMULA THAT INCORPORATES A MEASURE OF EDUCATIONAL NEED

Chapter I of this report has pointed out some of the shortcomings of the present state aid formula, particularly the fact that it recognizes differences in fiscal need among districts, but not differences in educational need. In Chapters II and III we have shown that it is possible to gather certain socioeconomic data through the local schools, and that these data are highly predictive of average achievement in the local school. This chapter will show how such data could be incorporated into a state aid formula. We will first discuss some general considerations that are involved in designing or changing a state aid formula. Next, we will define a new unit of educational need, called the "Need WADA," or "NWADA". Third, we will show several ways in which this measure of need can be incorporated into a state aid formula. Finally, we will show the possible results of several such formulas on some selected school districts in New York State, and recommend one of these formulas.

Considerations in Designing a Formula

An understanding of the Diefendorf formula starts with a definition of the Aid Ratio in terms of the WADA (Weighted Average Daily Attendance). This Aid Ratio is defined as follows:

$$\text{Aid Ratio} = 1 - .51 \times \frac{\text{District Full Value per WADA}}{\text{State Average Full Value per WADA}} \qquad (1)$$

This aid ratio is the heart of the Diefendorf formula, for it determines the proportion of local expenditures that the state will share, with the fiscally poorer districts receiving a larger proportion of their expenditures from the state than do the richer districts. This Aid Ratio is limited to a maximum of .90 and a minimum of .36. These limitations are important, and should be kept in mind when considering any formula in which the Aid Ratio appears. The basic Diefendorf formula is then:

State Aid = WADA x Expenditures per WADA x Aid Ratio (2)

In this formula, "Expenditures per WADA" are currently limited to $760, and since the vast majority of New York school districts spend more than that amount we might simply say that the formula is :

State Aid = WADA x $760 x Aid Ratio (3)

However, to do so obscures an important part of the original intention of the formula, which was to stimulate local effort by rewarding it with increased state appropriations. This is an important concept, even though the present $760 ceiling limits its application. In our discussion, then, we shall usually use "Expenditures per WADA" (abbreviated "Exp/WADA") in the formula rather than $760. A more complete statement of the present state aid formula is:

State Aid = (WADA x Exp/WADA x Aid Ratio) + Size Correction Aid + Urban Aid (4)

Chapter I of this report has listed some of the faults of Size Correction Aid and Urban Aid. The formulas we will propose will not use these kinds of aid.

We see three basic considerations in any discussion of the incorporation of a measure of educational need into a state aid formula. The first is whether or not the Diefendorf formula should be entirely scrapped and a new one devised. Our discussion above has implied that we do not recommend this. Some method of distribution is needed that takes account of differing fiscal abilities of districts. For all of its faults, the Diefendorf formula does a better job of this than most state aid formulas. We will recommend changes in, or additions to, the Diefendorf formula, but we do not propose to throw it out entirely. This does not preclude changes in that formula to take better account of differences in fiscal ability. However, such recommendations are the job of the other consultants to the Conference Board. When we recommend a formula that incorporates the present Diefendorf formula we assume that improvements in that formula would automatically be incorporated into our formula.

A second general consideration is whether the aid based upon educational need should be general aid or categorical aid. The Diefendorf formula is a general aid formula. The money received by a school district through it may be used for any purpose for which the school district may legally spend money. The transportation aid formula provides categorical aid. Money received by a district through it may only be spent on transportation of students. The important criterion that distinguishes categorical from general aid is that categorical aid may only be spent for a certain class of expenditures, or that there are other restrictions on its use that do not apply to general aid. Whether or not the aid is incorporated into the Diefendorf formula is not pertinent to this consideration. It is possible to incorporate

8

categorical aid into the general aid formula, just as it is possible to have a separate formula distributing either general or categorical aid. This matter of the formula is a separate issue that will be taken up below.

There are pros and cons to distributing aid based on educational need either as general aid or as categorical aid. The advantage of categorical aid is that the aid is directed at the need, with accounting controls to insure that the money will be used only for the purposes for which it is designed. This makes it difficult for the money to get siphoned into general faculty salary increases or other district-wide uses. But this advantage can also be a disadvantage. Such a formula usually brings with it a new bureau in the State Department of Education to supervise its distribution and use. A bureaucracy develops with a long-term interest in the stability and expansion of the program. Such a bureaucracy can result in the perpetuation of a program that is unimaginative and that persists long after the needs that inspired it have changed. We feel strongly that general aid will merely increase overall district expenditure without directly attacking the problems of the educationally disadvantaged, and we therefore favor categorical aid. But we have a recommendation that may eliminate some of the disadvantages of categorical aid listed above. We recommend that any additional money (above the general aid from the Diefendorf formula) received by a district based on a formula incorporating a measure of educational need be spent in the individual schools of the district in proportion to their contribution to that measure of need. This would insure that the money would not be used for district-wide increases in teacher salaries or otherwise spread evenly throughout the district. Appropriate accounting controls would be needed to see that this money was indeed spent in the schools in these proportions. But we would not insist that the money could be spent on only a certain class of expenditures. Educators have been less than omniscient so far in discovering how best to educate the disadvantaged. The money received by these schools should stimulate innovation, and this means that aside from designating the target population there should be no strings attached to it. It might be used to employ specialists, to decrease class size, to develop new curricula, or to bus some children to other schools. With enabling legislation it might even be used for such unusual ideas as paying slum children for good work to help motivate them. This money should not, however, be thought of as a slush fund. As rapidly as possible a method should be developed for fairly comparing the actual achievement in individual schools with the achievement predicted by the prediction equation discussed in Chapter III. Those schools in which actual achievement is significantly better than predicted presumably have a superior educational program that is well suited to the needs of its students. There should be some kind of reward for the personnel of such a school who are responsible for the program. The existence of such a reward should stimulate the search for truly effective innovations, for they would be tied to better-than-expected performance. On the other hand, there will be schools in which the actual performance is substantially below what would be predicted by the prediction equation. Outside help should be mandated for these schools, to help them improve their programs. We believe such a system of rewards for the effective and extra help for the ineffective to be very important, but it is not within the scope of this study to detail just how this might be done.

A third general consideration in the development of a formula based upon educational need is whether it should be a separate formula or whether it should be part of the Diefendorf formula. We do not feel strongly about this, but suspect that the formula will be less subject to attack by special interest groups if it is incorporated as part of the general formula. The formulas we will suggest will be of this type.

A fourth consideration is whether the distribution based upon educational need should be a constant number of dollars per unit of need, or should vary with the fiscal ability of the district. The answer to this depends partly upon whether the state intends to pay for the entire cost of this program, or to share its cost with the districts. If the state intends to pay for the entire cost of the program, there is no need to take local fiscal ability into account, because there is no local contribution. But if a local contribution is expected (and particularly if it is expected that the local contribution will be a large proportion of the total expenditure) there is reason to adjust the amount of the state contribution to the fiscal ability of the district. We will develop prototype formulas showing both flat grants and variable grants.

The NWADA and Aid Ratios Based on It

Our first step in the development of a formula is to define a new unit for measuring need. We could call this a "Need WADA" or, for short, an "NWADA". As explained in Chapter I, one might define "need" in terms of the number of children in the school who fall below a certain standard in achievement. But to base a state aid formula on such a definition would provide a disincentive for improvement. It would reward inefficient instructional performance. Instead, as explained in Chapter III, we propose to base the formula upon a prediction of that number of children, with the prediction based upon the socioeconomic status of the children in the school. Our criterion variable in Chapter III was R + A, the percentage of children below standard in reading plus the percentage below standard in arithmetic. Our NWADA will be a measure of the predicted number of these children, so that the NWADA could be defined as follows, for a single school:

$$\text{NWADA} = \text{WADA} \times (\text{predicted proportion below in reading} + \text{predicted proportion below in arithmetic}) \qquad (5)$$

We suggested four possible equations for predicting R + A in Chapter III. Those equations had regression weights based upon data from our sample, but we stressed that the final decision on which equation to use, and the weights to use for each variable, should be determined from statewide data. Let us generalize the prediction equation as follows:

$$R + A = (a \times N) + (b \times PR) + (c \times BrHo) + (d \times MOB) + (e \times SCH) + f \qquad (6)$$

where the variables have the same definition as in Chapter III (see Table III), and a,b,c,d,e, and f are constants, determined through the multiple regression statistical procedure, which constitute the weightings for the variables.

This prediction equation will predict the percentage below standard in reading and arithmethic. To convert it to a proportion we divide by 100. The definition of NWADA for a particular school would then be:

$$\text{NWADA} = \text{WADA} \times \frac{(a \times N) + (b \times PR) + (c \times BrHo) + (d \times MOB) + (e \times SCH) + f}{100} \qquad (7)$$

Since this NWADA is a prediction of the number of children who are below standard in reading plus the number who are below standard in arithmetic, it is theoretically possible for the NWADA in a school with a WADA of 100 to vary from zero (a prediction of no pupils below standard in either reading or arithmetic) to 200 (a prediction of all pupils below standard in both reading and arithmetic). Actually, there is probably no school in the state with no students below standard in reading or arithmetic, just as there is probably no school all of whose pupils are below standard in both subjects. Based upon the data from our sample, it appears that the reasonable limits of NWADA for a school with a WADA of 100 are about 10 and 150. The NWADA, then, is our measure of need, and is central to the development of a state aid formula based on such a measure.

For the formulas we will suggest below we will also need to define two new aid ratios. The aid ratio in current use for distribution of general aid is shown in Formula (1) on the first page of this chapter. Let us henceforth refer to this as the General Aid Ratio. In a similar fashion we can define a Need Aid Ratio, as follows:

$$\text{Need Aid Ratio} = 1 - p \times \frac{\text{District full value per NWADA}}{\text{State average full value per NWADA}} \qquad (8)$$

The letter "p" in the formula stands for a proportion in the form of a decimal between zero and one, equivalent to the .51 in the General Aid Ratio. The value of p could be set by the state depending upon the extent to which the state wished to share in the cost of a program based upon educational need. If the value of p were set at less than .51, the state would pay a greater share of the costs of this program than of the general costs of the districts, and the local share would be less. The reverse would be true if the value of p were set at more than .51. It would probably be desirable to set maximum and minimum limits on the value of the Need Aid Ratio, just as there are on the General Aid Ratio, but we are not recommending that these necessarily be set at the present 90% and 36%.

The second new aid ratio we will call the Combination Aid Ratio, and we define it as follows:

$$\text{Combination Aid Ratio} = 1 - .51 \times \frac{\text{District full value per (WADA + NWADA)}}{\text{State average full value per (WADA + NWADA)}} \quad (9)$$

The uses of these two new aid ratios will be illustrated as we develop some possible state aid formulas below.

Development of State Aid Formulas

We see the possible ways of distributing state aid based on educational need as being the following:

a. A grant requiring no participation by the local district. This could be in terms of a set number of dollars per NWADA. A formula that would do this would look like the following (the $y stands for an arbitrary number of dollars):

$$\text{State Aid} = (\text{WADA} \times \text{Exp/WADA} \times \text{General Aid Ratio}) + (\text{NWADA} \times \$y) \quad (10)$$

This formula would give each district a set number of dollars per NWADA in addition to what it received in general aid. On the other hand, the amount of the grant could be made to depend upon the actual local expenditures for the target group (but with a set maximum, of course). There would still be no local participation required, and therefore the amount of the grant would not depend upon the fiscal ability of the district. This formula would be as follows:

State Aid = (WADA x Exp/WADA x General Aid Ratio) + (NWADA x Exp/NWADA) (11)

b. A variable grant with local participation in which the state provides educational need aid according to the same ratio based on fiscal ability that is currently used. This could either be based on a fixed amount per NWADA, as follows:

State Aid = (WADA x Exp/WADA x General Aid Ratio)
+ (NWADA x $y x General Aid Ratio) (12)

or it could be based on actual expenditures, with a fixed maximum:

State Aid = (WADA x Exp/WADA x General Aid Ratio)
+ (NWADA x Exp/NWADA x General Aid Ratio) (13)

c. A variable grant with local participation where the educational need aid is provided according to the same aid ratio as general aid, but this aid ratio itself reflects the measure of educational need. This formula also might provide a set dollar amount per NWADA:

State Aid = (WADA x Exp/WADA x Combination Ratio)
+ (NWADA x $y x Combination Aid Ratio) (14)

or it might be based on actual expenditures with a fixed maximum:

$$\text{State Aid} = (\text{WADA} \times \text{Exp}/\text{WADA} \times \text{Combination Aid Ratio}) + (\text{NWADA} \times \text{Exp}/\text{NWADA} \times \text{Combination Aid Ratio}) \quad (15)$$

d. A variable grant with local participation in which the educational need aid is granted according to a different aid ratio than is the general aid. Again, the formula could be keyed to a set dollar amount per WADA or to a share of actual expenditures:

$$\text{State Aid} = (\text{WADA} \times \text{Exp}/\text{WADA} \times \text{General Aid Ratio}) + (\text{NWADA} \times \$y \times \text{Need Aid Ratio}) \quad (16)$$

$$\text{State Aid} = (\text{WADA} \times \text{Exp}/\text{WADA} \times \text{General Aid Ratio}) + (\text{NWADA} \times \text{Exp}/\text{NWADA} \times \text{Need Aid Ratio}) \quad (17)$$

We may begin to analyze these possibilities by noting that formulas (10), (12), (14), and (16) differ, respectively, from formulas (11), (13), (15), and (17) by the fact that the former are based on a set number of dollars per NWADA, while the latter are based on full or partial reimbursement for actual expenditures (up to a set maximum number of dollars per NWADA). It is probable that in practice there would be little difference between these two methods. The experience with the Diefendorf formula has been that the maximum is set so low that the vast majority of the state's districts spend more than this. Thus the incentive to increase local effort that was originally intended in the formula no longer exists for most districts. For all but a few districts in the state, the formula has $760 substituted for

Exp/WADA. We have no reason to believe that it would be otherwise with the educational need portion of the formula. Another difficulty with those formulas that are based on a sharing of actual expenditures is that they tend to favor the financially most able districts, and often these are the districts with the least in the way of educational need. Finally, when a formula consists of two portions, each based on actual expenditures, and these expenditures may be in the same categories, there is room for possible confusion and manipulation in accounting. For all of these reasons, we favor a formula that provides a fixed number of dollars per NWADA, subject perhaps to an aid ratio, but not to a cost-sharing arrangement.

This leaves us with formulas (10), (12), (14), and (16). Which of these is best is difficult to say without more knowledge than we currently have of the NWADA for each district in the state. Formula (10) is suitable if the number of dollars per NWADA is set at a figure that might be expected to cover all or most of the costs of the program. Formula (12) would distribute the need according to the same aid ratio as is currently used, and there may be some slight benefit to the simplicity of this arrangement. Formulas (14) and (16) would both tend to give more emphasis to the districts with the greatest educational need because the NWADA is used in two places in the formula.

Estimates of Effects of the Formulas

In an attempt to get some idea of the effects of the various formulas we have chosen some sample districts and attempted to make some estimates of how they might fare under sample formulas of each of the four types. Table XI gives estimated data for each of these districts. This is not intended to be a representative sample of districts. They were chosen for their interest and to illustrate the formulas. They consist of the Big Six city districts, three wealthy districts, three average districts, and three poor districts. There is a most important reservation to be kept in mind in reading the Table. We have been forced to estimate the NWADA for the district without knowing the socio-economic data to calculate it according to a formula such as that given earlier in this chapter. Remembering that the NWADA is a prediction of the number of children in a district below average in reading plus the number below average in arithmetic, we have instead estimated the NWADA by using the actual number below in reading plus the actual number below in arithmetic. If the schools in the district are performing in a manner that is typical of the state, this estimate may be fairly close to the NWADA computed from a prediction equation. But if the schools in the district are more ineffective than is expected, the NWADA estimated in our Table would be more than the actual NWADA. The figures of Need Aid given in the Table for that district would be higher than they would be if a program were actually in effect using that formula. The converse would be true if a district had unusually effective schools. For this important reason, the figures in the Table should only be looked upon as illustrative, not definitive. Columns 1 to 4 of the Table give actual data for the district on percent below standard in reading and in arithmetic, district WADA, and district full value per WADA (abbreviated FV/WADA). Column 5 gives our estimate of NWADA based on the information in the first three columns. Columns 6 and 7 give estimates of FV/NWADA and FV/(WADA + NWADA), which are necessary to calculate the aid ratios. Columns 8 to

TABLE XI

ESTIMATES OF STATE AID DISTRIBUTED BY VARIOUS FORMULAS TO A SAMPLE OF DISTRICTS

District	1967-68 (3rd Grade) % Below Standard in Reading[1] (Col. 1)	Arithmetic[1] (Col. 2)	1967-68 WADA[2] (Col.3)	1967-68 FV/WADA[2] (Col. 4)	1967-68 Est. NWADA[3] (Col. 5)	1967-68 Est. FV/NWADA[4] (Col.6)	1967-68 Est. FV/(WADA + NWADA)[5] (Col. 7)	General Aid Ratio[6] (Col. 8)	Need Aid Ratio[7] (Col. 9)	Combination Aid Ratio[8] (Col. 10)	1967-68 Basic Entitlement[9] (Col. 11)	1968-69 Size and Urban Aid[2] (Col. 12)	Need Aid Entitlement[10] Formula (18)[11] (Col. 13)	Formula (19) (Col. 14)	Formula (20) (Col. 15)	Formula (2 (Col. 16)
N. Y. City	46.5	50.8	1,025,000	44,300	997,000	45,500	22,500	.23(.36)	.60	.49	$280,440	$111,756	$199,400	$144,000	$238,000	$239,000
Albany	33.3	22.1	11,900	50,317	6,600	90,700	32,400	.19(.36)	.21(.36)	.19(.36)	3,260	882	1,320	950	950	950
Buffalo	39.8	29.5	72,200	28,446	50,000	41,100	16,800	.54	.64	.58	29,630	7,825	10,000	10,800	13,800	12,800
Rochester	33.5	49.2	44,000	41,086	36,400	49,700	22,500	.33(.36)	.57	.44	12,040	3,150	7,280	5,240	9,080	8,300
Syracuse	26.2	19.1	30,000	32,937	13,600	72,700	22,700	.47	.36	.43	10,720	2,533	2,720	2,560	1,420	2,010
Yonkers	22.4	20.3	29,800	43,226	12,700	101,400	30,300	.30(.36)	.11(.36)	.24(.36)	8,150	2,172	2,540	1,830	1,830	1,830
Webb 1	11.1	7.4	456	79,530	84	427,000	67,000	.00(.36)	.00(.36)	.00(.36)	125	0	17	12	12	12
Colchester 1	23.4	10.6	506	102,188	172	301,000	76,300	.00(.36)	.00(.36)	.00(.36)	138	0	34	25	25	25
Lackawanna	15.8	11.6	5,750	70,898	1,580	260,000	55,700	.00(.36)	.00(.36)	.00(.36)	1,570	36	314	228	228	228
Elmira	19.8	10.5	14,220	18,820	4,310	62,100	14,400	.70	.46	.64	7,570	407	862	1,210	450	793
East Greenbush	3.3	2.5	4,840	18,940	281	327,000	17,900	.69	.00(.36)	.55	2,540	79	56	78	-455	40
Huntington 10	12.2	5.5	13,170	19,216	2,330	109,000	16,300	.69	.05(.36)	.59	6,910	243	466	643	-455	336
Russell 1	22.2	5.5	650	4,754	180	17,200	3,720	.92(.90)	.85	.91(.90)	445	39	36	64	64	61
Hannibal 1	48.8	41.4	1,720	6,868	1,550	7,620	3,610	.89	.93(.90)	.91(.90)	1,160	96	310	552	574	558
Peru	11.9	4.9	3,600	7,631	605	45,800	6,540	.88	.60	.84	2,410	76	120	211	90	144

(All dollar amounts are in thousands of dollars)

NOTES TO TABLE XI

1. Data furnished by Bureau of Pupil Testing and Advisory Services, N.Y. State Education Department.
2. Data furnished by Division of Educational Finance, N.Y. State Education Department.
3. Estimated as follows: $\text{NWADA} = \text{WADA} \times \frac{\text{READ} + \text{ARITH}}{100}$. This assumes that NWADA as derived by the formula using socioeconomic data will be identical to this estimate using achievement test results. This limitation of these estimates should be clearly understood. See text for further comments.
4. Estimated as follows: $\text{FV/NWADA} = \text{FV/WADA} \times \frac{\text{WADA}}{\text{NWADA}}$.
5. Estimated as follows: $\text{FV/(WADA + NWADA)} = \text{FV/WADA} \times \frac{\text{WADA}}{\text{WADA + NWADA}}$.
6. $\text{General Aid Ratio} = 1 - .51 \times \frac{\text{District FV/WADA}}{\text{State Avg. FV/WADA}}$. State average FV/WADA was $31,500.
7. $\text{Need Aid Ratio} = 1 - .51 \times \frac{\text{District FV/NWADA}}{\text{State Avg. FV/NWADA}}$. State average FV/NWADA was calculated at $58,333 on the assumption that total state NWADA is 54% of total state WADA.
8. $\text{Combination Aid Ratio} = 1 - .51 \times \frac{\text{District FV/(WADA + NWADA)}}{\text{State Avg. FV/(WADA + NWADA)}}$. State average FV/(WADA + NWADA) was calculated at $20,454 on the assumption that total state NWADA is 54% of total state WADA.
9. Calculated from the formula WADA x General Aid Ratio x $760.
10. Need Aid entitlement is total district entitlement under a particular formula less its basic entitlement.
11. See text for Formulas (18) to (21).

10 give calculated aid ratios of the three types we have discussed in this chapter. Column 11 gives the entitlement under the basic Diefendorf formula using the WADA from Column 3 and assuming that each of the districts has expenditures per WADA that are at least $760. The amounts in this column exclude size correction and urban aid. Column 12 gives the actual entitlement of the district in size correction and urban aid for 1968-69. This would be excluded from any of the formulas we are recommending, and it is shown here for comparison with the amount of Need Aid the district would receive under the various formulas. Columns 13 to 16 show the amount of Need Aid the district would receive under each of four different formulas. Need Aid is defined as the difference between the total received under the formula and the amount of the basic entitlement under the Diefendorf formula. The four formulas are based on Formulas (10), (12), (14), and (16). For use in connection with formula (16) we have arbitrarily set "p" in the Need Aid Ratio to .51, the same as it is in the General Aid Ratio. We have arbitrarily set the dollar amounts at $200 for the first formula and at $400 for the other three. Thus, the four formulas used in the Table, representing the four different types we have discussed are:

$$\text{State Aid} = (\text{WADA} \times \$760 \times \text{General Aid Ratio}) + (\text{NWADA} \times \$200) \tag{18}$$

$$\begin{aligned}\text{State Aid} = {} & (\text{WADA} \times \$760 \times \text{General Aid Ratio}) \\ & + (\text{NWADA} \times \$400 \times \text{General Aid Ratio})\end{aligned} \tag{19}$$

$$\begin{aligned}\text{State Aid} = {} & (\text{WADA} \times \$760 \times \text{Combination Aid Ratio}) \\ & + (\text{NWADA} \times \$400 \times \text{Combination Aid Ratio})\end{aligned} \tag{20}$$

State Aid = (WADA x \$760 x General Aid Ratio)

+ (NWADA x \$400 x Need Aid Ratio) (21)

It would have been nice to use formulas with parameters such that they each distributed the same amount of money statewide as is now distributed by size correction and urban aid. The comparisons among them would then have been more meaningful. However, we have no way of estimating these parameters until statewide data are obtained. But, we can discern some effects. Formula (18) has something for everyone, including the very wealthy districts, which may get ten times as much as they now get in size correction and urban aid. It appears, on the basis of this small and unrepresentative sample, to do no better job of directing state aid to areas of educational need than do the present size correction and urban aid. Formula (19) is apparently even less redistributive. On the other hand, Formula (20) is so redistributive that some districts would not only lose all of their present size correction andurban aid, but would also lose some of their basic entitlement. This is because the Combination Aid Ratio contains a measure of need and applies both to the basic entitlement portion of the formula and to the need portion.

On balance, it appears to us that a formula of the form of Formula (21) may be the best possibility. We reject Formula (20) because it interferes with the general aid entitlement of the district (it is the only one of the four formulas that does this, as can be seen by comparing the first part of each formula with the basic Diefendorf formula). Formula (18) can be shown to be a special case of Formula (21), with the value of "p" in the Need Aid Ratio set equal to zero. Thus Formula (21) is a more general formula than Formula (18). And Formula (21) is more flexible than Formula (19) because both the number of dollars per NWADA and the proportion used in the Need Aid Ratio may be set by the state. We recommend Formula (21) and repeat it here in its most general form:

$$\text{State Aid} = \text{WADA} \times \text{Exp/WADA} \times \left(1 - .51 \times \frac{\text{District Full Value per WADA}}{\text{State Avg. Full Value per WADA}}\right)$$

$$+ \ \text{NWADA} \times \$y \times \left(1 - p \times \frac{\text{District Full Value per NWADA}}{\text{State Avg. Full Value per NWADA}}\right)$$

The "$y" is an amount to be based upon the amount the state intends to distribute in Need Aid. The "p" is a proportion that determines the aid ratio for the average district. By properly setting these two parameters the state can devise a formula that not only provides adequately for educational need, but makes political and common sense as well. The first part of this formula is just the basic Diefendorf formula, and constitutes general aid. The second part of the formula can be designated Need Aid, and it is a district's entitlement under this part of the formula that is to be spent in the individual schools of the district in proportion to each school's NWADA.

CHAPTER V

SUMMARY AND RECOMMENDATIONS

The children of poor families in New York State, as in America at large, obtain less and poorer education than do the children of other families. Even in those rare districts where the schools serving children of poor families provide facilities and faculties equal to those in other schools, the children do not achieve as well because of their impoverished backgrounds.

Through an extensive review of the literature we have confirmed the fact that educational achievement is highly correlated with socioeconomic status. We assert that persistent differences in average educational achievement among school districts or individual schools are indications of educational needs that are not being met. We do not allege that there should be no differences in educational achievement among individuals, for there are wide individual differences in intellectual ability. But we believe that it is the duty of the school to attempt to remedy those educational deficiencies that are imposed upon the child by his environment, and we know that this will cost extra money.

New York's original Diefendorf formula recognized that school districts vary in their local ability to finance education, by making the amount of state aid per pupil dependent upon a measure of local fiscal ability. However, it did not adequately recognize the great differences in the cost of providing a proper education for different students. It allowed 25%

more money per student for high school students. But it made no specific recognition of other kinds of students with special needs: the mentally retarded, the physically handicapped, the economically disadvantaged, and others. In other words, the Diefendorf formula recognized differences in fiscal need, but not in educational need. It is the purpose of this report to recommend changes that will appropriately recognize that there are differences in educational need among districts. We will present a way of measuring those differences and of using that measure in a distribution formula.

Since there had been special (and high cost) programs for many years before the Diefendorf formula for mentally retarded and physically handicapped children, it is clear that the framers of that legislation could not have completely ignored these areas in their deliberations. Instead, the implicit assumption was made either that whatever special needs exist are spread rather evenly among school districts, or that these disabilities are positively correlated with district wealth, so that special provision for them in the formula was unnecessary. Many recent studies have shown this to be an untenable assumption. Not only are the various disabilities that contribute to educational retardation unevenly spread among the school districts, but they are closely associated with the socioeconomic status of the families whose children attend schools in the district. This has caused an increasing need for educational services in those school districts serving a clientele composed of the lower socioeconomic groups at the same time that there has been a decreasing fiscal ability in many of these same districts. The resulting budget strain has forced some revisions in the Diefendorf formula. These changes (the "size corrections" and the "urban aid") have been applied piecemeal as budget pressures stimulated political pressures. The time has now come to make some basic changes in the formula.

We have undertaken a study of a representative sample of elementary schools in New York State. We have utilized certain socioeconomic measures that can readily be gathered by the local school authorities. We have correlated these measures with measures of educational achievement of the pupils in the school. Using the seven measures on which we collected data, and using the individual school as the unit of analysis, we can predict 75% of the variation in educational achievement among schools without knowing anything about the schools' instructional programs. Using only four of those measures, we can predict over 71% of the variation in achievement. The four are 1) the percentage of Negro and Puerto Rican students in the school, 2) the percentage of children from broken homes, 3) the average number of different schools the children in the school have attended in the last three years, and 4) the average number of years of schooling of the parents of the children in the school. We use these factors, with statistically determined weightings, in an equation predicts the percentage of children in the school who score below an acceptable level on the state-administered achievement tests in reading and arithmetic. These achievement tests do not measure all of the kinds of things the schools attempt to teach, but they are objective measures of some very important learning areas, and the results of them would probably correlate well with measures of achievement in other areas.

The limits of time and money have made it impossible to obtain a completely representative sample, although the remarkable predictive powers of the variables in our sample encourage us to believe that the sample is substantially representative.

One could, of course, simply use results of the achievement tests as a measure of educational need and distribute state money on that basis. In fact, the results of these tests are currently being used as one factor in the "urban aid" portion of the distribution formula. But this procedure leaves open the possibility of a reward for poor performance on the part of the teachers and administrators. Instead, we propose to predict what educational achievement would be for the type of children in the school by the use of a formula such as the sample one we have developed, and distribute state aid on that basis. Consideration should then be given to some method of rewarding those schools whose students achieve significantly better than predicted, and of giving outside assistance to those schools whose students achieve significantly more poorly than predicted, but it is not within the scope of this report to detail how that might be done.

The individual school is the unit of our analysis. The data we collected is based upon the individual students who actually attend that school rather than any measures of the socioeconomic status of the school's neighborhood or attendance area. Our measure is thus an accurate reflection of those who are actually being educated in the school, regardless of where they live.

We chose the school as the unit of analysis for several reasons:

a. The individual child is too small a unit of analysis. As previously mentioned, there will always be large variations in individual achievement because of differences in innate ability.

b. The school district is too large a unit. In most districts there are schools with a concentration of disadvantaged children and other schools with a concentration of advantaged children. When using district averages many of these schools cancel one another out, leaving a composite figure that hides real educational problems.

c. School district consolidation or decentralization would be unlikely to affect a measure based on the individual school.

d. The school is the natural unit of the educational organization, and is therefore the level at which we may best look for educational improvements.

This last reason is also the rationale for our later recommendation that the additional money that a district would receive through a state aid formula based upon a measure of educational need should be spent in the individual schools of the district in proportion to their contribution to that measure of need, rather than spreading the additional money evenly over the district. However, if this money is to stimulate innovation there should be no restrictions on how it is used in these schools other than the general legal restrictions to which all school districts are subject.

Chapter IV details several different ways of developing a formula based upon educational need. All of them depend upon the definition of a new need measure based on the WADA (Weighted Average Daily Attendance). We call this new measure the "Need WADA", or the "NWADA" for short. A school's NWADA is a prediction (based upon a socioeconomic formula) of the number of children in the school who will achieve below standard in reading plus the number who will achieve below standard in arithmetic. Technically, the NWADA may be defined as follows:

$$\text{NWADA} = \text{WADA} \times \frac{(a \times N) + (b \times PR) + (c \times BrHo) + (d \times MOB) + (e \times SCH) + f}{100}$$

where N = percentage of Negroes in the school

PR = percentage of Puerto Ricans in the school

BrHo = percentage of children in the school from broken homes

MOB = mobility of children (average number of schools attended by the school's pupils in the last three years)

SCH = average number of years of schooling of the parents of children in the school

a,b,c,d,e, and f are constants developed through statistical treatment of the data.

Remembering that the NWADA is a prediction of the number of children who will be below standard in reading plus the number who will be below standard in arithmetic, we can see that, for a school with a WADA of 100, the lowest possible NWADA would be 0, and the highest possible NWADA would be 200 (equivalent to all children below standard in both reading and arithmetic.) Based on our data, the practical limits of NWADA for a school with a WADA of 100 are about 10 and 150.

We recognize that even in the best schools there will be some children who do not achieve up to standard. It could be argued that such a school should not receive aid on the basis of special educational need; that only schools with more than average need should receive this aid. But there is a flaw in such reasoning. In such a case, a district with one disadvantaged school and one advantaged one would receive aid for the disadvantaged school. If the district then changed attendance boundaries or bused to achieve racial balance, it might have two average schools, neither of which would receive aid, yet it would still have the same pupils with the same problems. Thus, a plan based on aid only to those schools worse than average will work against achievement of racial balance. For this reason we recommend aid to all schools on the basis of of their NWADA. This is not qualitatively different from the present situation where all schools receive size correction aid.

The present Diefendorf formula is:

$$\text{State Aid} = \text{WADA} \times \text{Exp/WADA} \times \left(1 - .51 \times \frac{\text{District FV/WADA}}{\text{State Avg. FV/WADA}}\right) + \text{Size Correction Aid} + \text{Urban Aid}$$

We have suggested four types of formulas incorporating the NWADA which could be used instead of the present formula. All of them continue the basic features of the Diefendorf formula while deleting size correction aid and urban aid. We have recommended one of these types of formula, which looks like this:

$$\text{State Aid} = \text{WADA} \times \text{Exp/WADA} \times \left(1 - .51 \times \frac{\text{District FV/WADA}}{\text{State Avg. FV/WADA}}\right) + \text{NWADA} \times \$y \times \left(1 - p \times \frac{\text{District FV/NWADA}}{\text{State Avg. FV/NWADA}}\right)$$

In the formula, Exp/WADA means "Expenditures per WADA," and "FV/WADA" means "Full Value per WADA." The "$y" in the formula is an arbitrary number of dollars, and the "p" is an arbitrary decimal between zero and one. Both may be adjusted to fit the amount of state money available for this program and the desired distribution of it depending upon district wealth. The effect of such a formula upon each of the districts of the state, and the state-wide costs, cannot be predicted with accuracy until complete data with which to compute the NWADA for each district are obtained, although Chapter IV gives some estimates for a few districts.

In addition, it would be unwise to base a state aid formula upon the weightings we have developed from our sample. Instead, the state should mandate the gathering of data by all districts on variables we have identified. A formula would then be developed based upon the formula above, using the solid data base of the entire state instead of the data of a sample. This formula would go into effect in the fiscal year following that in which the data were gathered.

Our specific recommendations may now be summarized as follows:

1. The Legislature should state its intention that a state aid formula incorporating a measure of educational need shall be placed in effect as soon as complete state-wide data can be gathered to establish the measure.

2. The measure of educational need shall incorporate the variables we have shown to be excellent predictors of educational achievement.

3. The extra money granted to school districts by this formula (above that given by the basic Diefendorf formula) shall be used in those particular schools where the educational needs are greatest.

These three major recommendations should be implemented as follows:

Step One: As soon as authorized by the Legislature, each public school in the state shall be required to gather the following information about its students:

a. Percentage of Negro students.

b. Percentage of Puerto Rican students.

c. Percentage of students from homes where one or both parents are missing.

d. Average number of schools attended by the school's pupils in the last three years.

e. Average years of schooling of the parents of the school's pupils.

f. Percentage of pupils in tested grade levels in the school who score below the fourth stanine in reading in the state achievement test.

g. Percentage of pupils in tested grade levels in the school who score below the fourth stanine in arithmetic in the state achievement test.

Step Two: Use the data so gathered, by use of the multiple regression statistical technique, to define the NWADA (a prediction of the number of students in a school below standard in reading plus those below standard in arithmetic) by means of the prediction equation given earlier in this statement.

Step Three: Develop a formula for distribution of state aid using the NWADA as an ingredient of the formula, as recommended earlier in this summary. Such a distribution would replace the present size correction aid and urban aid.

Step Four: Provide a distribution method that insures that this additional aid based on educational need is spent in the individual schools in proportion to their NWADA, instead of being spread evenly over the school district, but leave districts and their schools a free hand, within usual legal limits, in determining how to spend the money.

Step Five: Provide that this new formula shall go into effect in the fiscal year following that in which the data are collected.

Step Six: Consider the possibility, after some experience is gathered with this system, of providing a method of rewarding those schools where the achievement is significantly better than predicted by the formula, and of mandating outside help for those schools where achievement is significantly worse than predicted.

APPENDIX A

TEACHERS COLLEGE, COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10027

MEASUREMENT OF EDUCATIONAL NEED PROJECT

(Date)

(Superintendent's Name)
(Address)

Dear :

I am engaged in a study commissioned by the New York State Educational Conference Board which will hopefully result in a recommendation to the next Legislature for revision in the state aid formula to better recognize the special financial needs of those school districts with culturally disadvantaged children. The Conference Board which is supported by associations of teachers, administrators, school boards and parents, has been very effective in the past in influencing the Legislature to increase the levels of state aid.

In order to support such a recommendation, it is necessary to gather supporting data. (Name of School) in your district has been chosen by a random sampling technique to supply some data, and I am earnestly requesting your cooperation in supplying it. This school, along with about 50 others, will form a sample that is representative of the entire state.

Within the school, the sample will consist of twenty students from the fourth grade(s) of the school. For each of these students I will need information on seven items:

1. State achievement test scores in the third grade.
2. Number of schools the child has attended in the last three years.
3. Number of rooms per person in the student's dwelling.
4. The student's racial or ethnic status.
5. Whether the student is on welfare.
6. How many of his parents the student lives with.
7. The number of years of schooling of his parents.

Some of the data will be available in the school's records; some will have to be obtained from the pupils themselves or their parents. Complete anonymity of students is guaranteed. We will use a coding system that will make it unnecessary for you to furnish us the names of the students.

It should not take much time to gather this data for twenty students. On a trial run in a ghetto school it took about five hours. However, we realize that this request intrudes on the time of busy people. The project is willing to pay for extra clerical time necessary for the data gathering.

I sincerely hope you are willing to cooperate in this project by forwarding this letter and the attached material promptly to the principal of the selected school. Prompt action is necessary if we are to prepare timely recommendations for the Legislature.

The rest of this letter is intended primarily for the principal.

Attached are materials necessary in collecting the data for this project, including:

1. Detailed instructions.
2. Form on which data are to be entered.
3. Reimbursement form for extra clerical time involved.
4. Return Envelope.

I believe you will find that the data to be gathered will be relatively easy to obtain if someone who is familiar with the students and their records (such as the school secretary) does the job. Your cooperation in this project is not only a professional service, but could result in a change in the state aid formula that would benefit your school.

Prompt completion of the data gathering is essential to the project. I hope that you will be able to return the data form to me by the date indicated on the instructions. If you have any questions, do not hesitate to telephone collect either Walter I. Garms, 212-870-4891, or Mark C. Smith, 212-870-4687.

Sincerely,

Walter I. Garms
Project Director

APPENDIX B

Teachers College, Columbia University
Measurement of Educational Need Project

Instructions for Collection of Data

The project for which you are gathering data may be of vital importance to the future of public education in New York State. The results of this project will probably have a large influence on changing the method of distributing state aid to reflect need for educational services. Your cooperation in gathering these data promptly and accurately is earnestly solicited. If possible, the information should be gathered, entered on the attached sheet, and returned by (date) . The detailed instructions follow:

1. You are going to collect data for a random sample of fourth graders in your school. We need complete data for 20 students. Using class lists or any other convenient method, choose every (nth) student. This will give you more than 20, to allow for not being able to get complete information on some. As soon as you have complete information on 20 students you are finished. It is important for sampling purposes that you use only the names selected by the above process, rather than choosing those for whom it is most convenient to get data.

2. On the attached data sheet list the students in Column A in some way that will allow you to identify them. If you wish to list their names you may, but we are not interested in the names. You might list the child's room number and his initials, for example. The main thing is that if we have a question about a particular item of data and must call you about it, we must be able to identify it so that you will know which child we are asking about.

3. Each of the pupils on the list, if he was in a New York State school last year, took a standardized achievement test administered under state auspices. The test results will normally be found in the pupil's permanent record. They will either be recorded as a percentile rank (generally a number between 10 and 100), or as a stanine score (a number from 1 to 9). Record in Column B whichever is available, but if both are available, report percentile rank only. Record both reading and arithmetic scores. Indicate at the top of column B whether you are recording percentiles or stanines.

4. Enter in Column C one of the following letters:

   N if the student is a Negro.

   P if the student is Puerto Rican.

   F if the student is not Puerto Rican, but the language usually spoken at home is not English.

   O if the student fits none of the above categories.

   There is often some question as to whether or not a student is Negro or Puerto Rican. For the purposes of this study, you decide by what he is commonly considered to be. The opinion of the principal or the teacher should be solicited in case of doubt.

5. Some of the rest of the information may be on the school records, but some of it will have to be gathered by direct contact with parents. Consult with your principal on the best way of gathering each item of data. Some of the questions may be particularly sensitive in some communities, and discretion is essential. A suggested form for a parent questionnaire is enclosed, which may be changed to suit local needs. Note that it does not ask for number of parents living with the child. You are to infer that from the number of parents who sign the form, or from other information available to you.

6. Enter in Column D the number of parents currently living with the child (either 2,1, or 0). A step-parent is counted as a parent, but a foster parent is not.

7. Enter in Column E either

   Yes if the child's family is receiving Aid for Dependent Children.

   No if they are not receiving this aid.

   You can get fairly reliable school data on this from participation in the Free Lunch program.

8. Enter in Column F the number of the highest grade in school completed by each parent or step-parent who is currently living with the child. For example, enter in the column for Father:

   7 if he completed the 7th grade

   10 if he finished two years of high school

   12 if he completed high school

   16 if he graduated from college

   N if that parent is not currently living with the child

9. Enter in Column G the number of rooms in the student's dwelling. Count only such rooms as kitchen, living room, dining room, bedrooms, family room, etc. The following kinds of rooms should not be counted: bathrooms, hallways, garage, etc.

10. Enter in Column H the total number of people who actually live in the dwelling unit, including unrelated persons such as a boarder. Do not include temporary visitors.

11. Enter in Column I the total number of different schools the student has attended during the last three years (1966-67, 1967-68, and 1968-69). If he attended one school, transferred to another school, and came back to the first, you would record as a change of schools the change that occurs as a result of grade organization of the school (as, for example, when the students of a school with grades 1-3 transfer to a school with grades 4-6.)

12. Your comments on difficulties you encountered, directions that are not clear, and your suggestions for improvement are earnestly solicited. Put them on a separate sheet of paper and attach it to the data sheet.

13. If you have any questions that are not covered by these instructions, or if for any reason you cannot complete the job by the date given in Instruction 1, please call the following:

Walter I. Garms	870-4891
Mark C. Smith	870-4687

14. A self-addressed envelope has been enclosed for your use in returning the form.

APPENDIX C

TEACHERS COLLEGE, COLUMBIA UNIVERSITY

Measurement of Educational Need Project
Data Sheet

School Name ____________________

Address ____________________

Telephone (include area code) ____________________

School District ____________________

Principal's Name ____________________

Name of person preparing this form ____________________

When completed return this form to:

Professor Walter I. Garms
Box 74
Teachers College
525 West 120th St.
New York, N.Y. 10027

A	B		C	D	E	F		G	H	I
Pupil Identification (Instruction 2)	3rd Grade Achievement Scores (Instruction 3)		Race or Ethnic Status (Instruction 4)	Parents Living With Child (Instruction 6)	Welfare Status (Instruction 7)	Parents' Education (Instruction 8)		Rooms in Dwelling (Instruction 9)	Number Living in Dwelling (Instruction 10)	Schools Attend (Instr tion
	Reading	Arith				Father	Mother			

APPENDIX D

TEACHERS COLLEGE, COLUMBIA UNIVERSITY
MEASUREMENT OF EDUCATIONAL NEED PROJECT

Claim for Reimbursement

Name ______________________________

Address ______________________________

____

Social Security No. ______________________

Hours of extra clerical time ________________

Hourly rate $________________

Total reimbursement claimed $________________

Signed:

____

Claimant

____

Principal

APPENDIX E

RAW DATA FOR 45 SAMPLE SCHOOLS

SCHOOL NUMBER	READ	ARITH	N	PR	F	BrHo	Wlfr	O-C	MOB	SCH
New York City:										
1	0.0	5.0	0.0	0.0	5.0	15.0	0.0	20.0	1.35	12.7
2	5.0	15.0	0.0	0.0	0.0	10.0	0.0	20.0	1.05	12.4
3	42.1	47.4	26.7	26.7	0.0	26.7	26.7	31.6	2.16	10.4
4	31.6	31.6	94.7	5.3	0.0	36.8	42.1	52.6	1.21	10.7
5	15.0	25.0	0.0	0.0	0.0	0.0	0.0	5.0	1.15	13.3
6	61.1	61.1	44.4	55.6	0.0	66.7	50.0	33.3	1.38	7.9
7	10.0	5.0	0.0	0.0	0.0	5.0	0.0	5.0	1.25	13.6
8	47.1	41.2	11.7	41.2	0.0	11.7	11.7	70.6	1.29	8.3
9	55.0	60.0	20.0	45.0	0.0	20.0	35.0	45.0	1.85	11.2
10	37.5	37.5	37.5	37.5	0.0	62.5	56.2	62.5	1.81	9.5
Other Large and Medium Cities:										
11	60.0	35.0	85.0	0.0	0.0	50.0	30.0	55.0	1.35	9.3
12	30.0	10.0	50.0	0.0	0.0	15.0	25.0	50.0	1.45	8.6
13	5.0	5.0	0.0	0.0	0.0	5.0	0.0	10.0	1.40	12.6
14	30.0	30.0	5.0	0.0	0.0	20.0	15.0	20.0	1.35	9.5
15	55.0	25.0	0.0	0.0	0.0	25.0	40.0	30.0	1.20	10.8
16	41.2	17.6	35.2	5.9	0.0	35.2	23.5	23.5	1.23	10.4
17	40.0	20.0	15.0	0.0	0.0	15.0	5.0	10.0	1.40	12.7
18	25.0	10.0	35.0	0.0	0.0	10.0	0.0	10.0	1.60	13.2
Small Cities and Suburbs:										
19	17.6	5.9	0.0	0.0	0.0	23.5	0.0	5.9	1.23	12.9
20	40.0	20.0	0.0	0.0	0.0	15.0	5.0	25.0	1.20	11.5
21	15.0	0.0	0.0	0.0	0.0	0.0	0.0	10.0	1.15	13.2
22	30.0	0.0	0.0	0.0	0.0	10.0	5.0	20.0	1.15	11.9
23	25.0	0.0	0.0	0.0	0.0	6.3	0.0	0.0	1.20	13.9
24	20.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	1.10	12.7
25	0.0	0.0	0.0	0.0	0.0	5.0	5.0	0.0	1.05	12.5
26	20.0	10.0	0.0	0.0	0.0	5.0	0.0	10.0	1.25	12.4
27	5.6	0.0	0.0	0.0	0.0	5.6	0.0	22.2	1.10	12.8
28	10.0	0.0	0.0	5.0	0.0	0.0	0.0	45.0	1.50	9.9
29	20.0	5.0	0.0	0.0	0.0	5.0	0.0	5.0	1.30	14.0
30	10.0	0.0	0.0	0.0	10.0	5.0	0.0	10.0	1.30	14.7

APPENDIX E (CONT.)

SCHOOL NUMBER	READ	ARITH	N	PR	F	BrHo	Wlfr	O-C	MOB	SCH
Small Cities and Suburbs (Cont.):										
31	15.8	5.3	5.3	0.0	0.0	10.5	0.0	0.0	1.10	15.6
32	15.0	5.0	0.0	0.0	10.0	20.0	10.0	15.0	1.15	11.3
33	10.0	10.0	0.0	0.0	0.0	5.0	0.0	20.0	1.05	12.6
34	23.5	0.0	0.0	0.0	0.0	5.9	0.0	23.5	1.23	13.8
Rural:										
35	10.5	0.0	0.0	0.0	0.0	0.0	0.0	15.8	1.10	12.9
36	20.0	5.0	0.0	0.0	0.0	15.0	10.0	10.0	1.10	12.2
37	30.0	10.0	0.0	0.0	0.0	0.0	0.0	15.0	1.10	12.0
38	15.0	0.0	0.0	0.0	0.0	0.0	0.0	50.0	1.05	11.3
39	41.2	29.4	5.9	29.4	0.0	5.9	0.0	11.8	1.17	9.9
40	40.0	10.0	0.0	0.0	0.0	0.0	0.0	15.0	1.20	12.6
41	5.3	15.8	0.0	0.0	0.0	0.0	0.0	22.0	1.00	11.8
42	25.0	5.0	0.0	0.0	0.0	10.0	5.0	15.0	1.10	10.0
43	45.0	15.0	35.0	0.0	0.0	10.0	10.0	10.0	1.55	10.0
44	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	1.45	14.6
45	45.0	5.0	0.0	0.0	0.0	10.0	0.0	15.0	1.65	13.2

Note: See Table III in text for definitions of variables shown above.